Body Motion Cues Drive First Impressions: Consensus, Truth and the Origins of Personality Trait Judgements based on Targets' Whole-Body Motion.

THORESEN, JOHN,CHRISTIAN

Body Motion Cues Drive First Impressions: Consensus, Truth and the Origins of Personality Trait Judgements based on Targets' Whole-Body Motion.

John Christian Thoresen

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy.

Department of Psychology

Durham University

2012

John Christian Thoresen

Body motion cues drive first impressions.

Personality trait attribution is automatic, and first impressions can be lasting and lead to important social decisions. Research on how facial cues impact on person perception is plentiful, but less is known about how whole-body motion contributes to first impressions. This thesis presents results from experiments in which participants rated the traits of target individuals based solely on short, silent movie clips of those individuals performing actions or expressing emotions with their bodies, or simply walking. To isolate the contribution to trait attribution of body motion cues, the static form information of the body stimuli was degraded. Consensus at zero acquaintance is replicated throughout the thesis, as manifested by strong inter-rater agreement for all rating experiments and within all displayed behaviours, thus indicating that body motion may contain visual cues that drive trait impressions. Further experiments identified motion parameters that predict personality trait impressions, and an experimental paradigm showed that computational manipulation of motion data can indeed change observer judgements of computerised models based on human motion data. No accuracy was found in the trait judgements, in that there was no link between how a target was judged and this target individual's scores on a five-factor personality questionnaire. Underlying judgements driving personality trait impressions were found; impressions of emotions, attractiveness and masculinity appear to be intertwined with personality trait judgements. Finally, patterns in personality trait judgements based on body motion were consistent with findings from studies on face perception, reflecting a two-step judgement of a target person's intention and ability to cause harm. Differences were found depending on the display format of the stimuli, and interpretations for these discrepancies are offered. The thesis shows that people go beyond the information available to them when forming personality trait impressions of strangers, and offers evidence that changes in body motion may indeed have an impact on such trait impressions.

Table of Contents

DISCUSSION

List of Tables

List of Illustrations

Statement of copyright.

Acknowledgements:

First and foremost, I am grateful to my supervisors Anthony Atkinson and Quoc Vuong without whose help this thesis would not be reality. Tony, for going out of his way to offer support and guidance throughout the PhD process, and for his long hours to provide timely and thorough responses to all my nagging emails. And Quoc, for supplying so much support for analyses, programming and write-ups, not to mention his persistence during our fights over APA formatting rules.

My very special thanks go also to Vera Kempe, whose continued belief in me has brought me to where I am today.

I would also like to thank Lucy Ferrara, Elena Broggi, Lina Aimola, Ala Hola and of course the German crew, for providing moral support and for reminding us all that academics too can have fun. Further, it would be inappropriate not to acknowledge my many students, whom I thank for experiment participation, inquisitive enthusiasm and generally for offering a much-welcomed breath of fresh air to the research environment at Durham University’s Psychology Department.

Finally, I am grateful to the discoverers of to $C_8H_{10}N_4O_2$—a compound central to my thesis work,—to all my friends and family for always being there for me, and to Michael O’Leary for allowing me to visit said friends and family on a teaching assistant salary.

INTRODUCTION

Chapter I: General Introduction

At the heart of this thesis lies the simple observation that, whenever people meet, they judge one another. These impressions are formed quickly and with great ease. They can be lasting, irrespective of accuracy, and they can drive important social decisions, irrespective of the impressions' relevance to these social decisions.

Background

People are quick to form impressions of a person's traits, which are considered to be stable dispositions to certain behaviours (Allport, 1961; Asch, 1946; Bar, Neta, & Linz, 2006; van't Wout & Sanfey, 2008). Observers decide how trustworthy, competent or likeable they think a target individual is after only 100 ms exposure to a picture of the target individual's face (Willis & Todorov, 2006); perception of threatening traits can be done as fast as 39 ms (Bar, et al., 2006): when given more time to make decisions, ratings do not change significantly. Observers often agree with each other on these impressions: what is known as consensus at zero acquaintance (Albright, Kenny, & Malloy, 1988). Such agreement has above all been found for personality traits when observers see photographs of strangers (Albright, et al., 1997; Penton-Voak, Pound, Little, & Perrett, 2006).

It has been found that first impressions can be lasting (Berry, 1990; Kenny, Horner, Kashy, & Chu, 1992), and can affect important social decisions. For instance, voting choice has been found to be linked to trait impressions based on faces (Little, Burriss, Jones, & Roberts, 2007; Mattes, et al., 2009; Todorov, Mandisodza, Goren, & Hall, 2005) and motion cues from the upper body (Kramer, Arend, & Ward, 2010). Faces of politicians rated as competent and non-threatening by separate observers were likely to win elections (Mattes, et al., 2009). Implicit trustworthiness judgements based on photographs have also been found to cue strategic decision making in trust games, even when no relationship existed between perceived trustworthiness and likelihood of reciprocal behaviour (van't Wout & Sanfey, 2008).

If judgements are so readily made, it is natural to assume that they are made as early as possible. For instance, the decision to avoid an individual is more effective if the observed person is at a great distance. Consequently, it is plausible, or indeed probable, that separate nonverbal cues may individually drive trait impressions; for instance, people may make judgements based on information of a person's body if they are too far away, or it is too dark, to see their face. Presented with only static body information, observers agree in judgements of attractiveness (Fan, Dai, Liu, & Wu, 2005; Fan, Liu, Wu, & Dai, 2004), which has been shown to be correlated with judgements of extraversion when observers are unacquainted with the target individuals (Albright, et al., 1988; Kenny, et al., 1992). However, dynamic information from a body—available at great distances and in visually limited settings—is also arguably likely to affect first impressions.

There is indeed support for the idea that body motion cues alone drive person-perception. People are good at making judgements such as the sex of a target (Kozlowski & Cutting, 1977), or the identity of a target (Cutting & Kozlowski, 1977; Loula, Prasad, Harber, & Shiffrar, 2005) based on body motion. Exaggerating arm movements of a target improves recognition accuracy of individuals (Hill & Pollick, 2000) or actions (Pollick, Fidopiastis, & Braden, 2001), which adds support to the proposition that it is, in this case, the movement that drives the person perception. Some social judgements have also been found to converge using dynamic stimuli. For instance, Gunns, Johnston and Hudson (2002) showed that observers make reliable judgements of how vulnerable others are when they are exposed to dynamic cues alone, showing that observers may indeed select victims based on cues from bodily motion. The study used point-light displays of targets walking naturally—stimuli in which static information is reduced or even eliminated (Johansson, 1973)—and the authors further controlled for gender and physical strength. Short video clips of targets unknowingly being filmed also show consensus at zero acquaintance for extraversion ratings (Ambady & Rosenthal, 1993; Kenny, et al., 1992).

Many of the studies cited previously prompted observers for a response; such judgements may thus be separate from implicit impressions. Implicit impressions are those that are at play when observers know that someone is acting ‘out of character’ but you they are not quite able to state how or why (Uleman, Blader, & Todorov, 2005). However, implicit and explicit impressions are intertwined; and articulating an impression for the first time means it has moved from implicit to explicit (Uleman, et al., 2005). Implicit trustworthiness judgements from faces, as measured through an observer’s willingness to invest in someone, have been found to converge subjective, explicit, trustworthiness judgements from the same faces (van't Wout & Sanfey, 2008).

The propensity to automatically infer personality traits could be a necessary consequence of limited cognitive ability: people categorise others due to a need to reduce the complex social environment (Uleman & Moskowitz, 1994). For instance, there exists a bias in attributing intentions to actions even when observers are told the actors had no choice (E. E. Jones, 1979). Correctly inferring personality traits may have a reproductive advantage, such as detecting whether someone is likely to be a monogamous partner and a considerate parent. Indeed, some studies report a kernel of truth in trait attribution based on nonverbal stimuli (Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004; Penton-Voak, et al., 2006). However, research on accuracy in trait attribution is inconsistent, especially for dynamic nonverbal stimuli (Kenny, et al., 1992; Montepare & Zebrowitz-McArthur, 1988).

The observation that people automatically make personality judgements and the social consequences of such judgements—despite the limited accuracy—makes studies of person-perception all the more interesting. How one is perceived by others can have implications for many aspects of life, such as succeeding at sales or getting along with work colleagues—or indeed getting a job in the first place— and forming friendships or relationships.

Indeed, personality research and research on the *perception* of personality could be considered two different things. Personality research has been occupied by describing systematic differences between people. In its most basic form, such research may follow Allport's assumption that personality is about real intra-psychic processes: 'personality is the dynamic organization within the individual of those psychophysical systems that determine his unique adjustment to the environment' (1937, p. 48). However, psychologists interested in perception of personality may, in fact, ignore Allport's view altogether: personality is simply how people are perceived. As such, the real consequences of personality, if there is such a thing, are the outcomes of how people are perceived in different settings (Funder, 1995; Hogan, 1996; Swann, 1984).

If there indeed is such a thing as a structured inner aspect of a person, this might thus be treated differently from the communicated aspect of the self; that is, the impressions that this person gives off to observers. This view is central to this thesis; that is, in many instances it is how someone is *perceived* that is of importance. It matters little if a job candidate is apt at performing his or her duties if the recruitment team is not convinced of this during the interview.

Several nonverbal cues could be explored for trait impressions: these could be auditory or visual, static or dynamic; and could be based on the body or the face, not to mention external displays of jewellery, clothing, tattoos etc. This thesis focuses on body motion because it is a cue that is available to observers when faces cannot be seen (e.g., at great distances). Further, research on personality attribution has largely focused on faces, or static information from the body, but less is known about personality trait attribution based on dynamic body stimuli. This is surprising, considering that it has been shown that people use dynamic stimuli to recognise others, discriminate gender or recognise actions.

The main aims of this thesis are to explore the link between whole-body motion of a person and that person's perceived personality. We assess consensus at zero acquaintance; that is, inter-rater agreement in personality trait judgements based on the nonverbal stimuli. We also explore cues to this consensus through computerised motion analysis, subjective ratings, and the impact of the type of behaviour displayed (action or emotion).

Exploring gender effects of observers and targets is a field not extensively covered in this thesis. It is known that body motion is important for recognising gender (e.g., Troje, 2002) and that social stereotypes of gender may affect trait attribution (e.g., Williams, Satterwhite, & Best, 1999). Therefore, certain steps have been made to control for target gender throughout the thesis. However, no analyses have been made to look at effects of observer gender. The literature suggests a female advantage for distinguishing biological from non-biological motion, as well as recognising emotions from point-light walkers (Alaerts, Nackaerts, Meyns, Swinnen, & Wenderoth, 2011). However, studies of other forms of person perception have not reported gender differences; for instance, male and female observers agree with one another in their judgements of the physical attractiveness of male targets and female targets, respectively (Berscheid, Dion, Walster, & Walster, 1971).

In this introductory chapter, we present a brief description of the experimental method used in the thesis, including personality measurements and motion analyses. We then present an overview of the experimental chapters.

Experimental Method

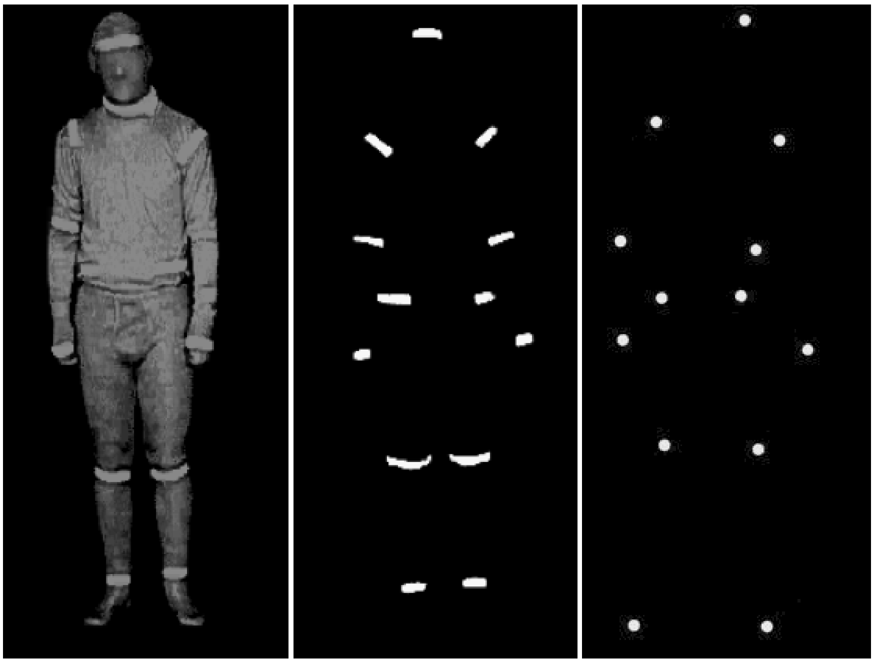
The Big Five. It is central to psychological researchers that whatever process we are interested in must be measurable. There is today widespread use of the five-factor model of personality; that is, a model consisting of five higher-order constructs, or traits, often agreed to carry the names Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness to experience (Costa & McCrae, 1988; Costa, Terracciano, & McCrae, 2001; Digman, 1990, 1997; McCrae, Costa, & Wiggins, 1996; McCrae & John, 1992). Costa and McCrae, developers of the NEO Five-Factor Inventory (FFI) (Costa & McCrae, 1992), argue that traits are, in fact, real units of personality originating in underlying neurophysiological structures (McCrae, et al., 1996). The five-factor model, often dubbed The Big Five, is derived primarily from questionnaires, but guided by the earlier research using a lexical approach (Allport & Odbert, 1936; Galton, 1884; McCrae & John, 1992). Its robustness is confirmed through the finding that almost identical patterns of personality traits are found across self-reports, spouse ratings and expert ratings (McCrae & Costa, 1989). External criteria also support the five-factor model: self- and peer- reported traits have been found to predict social status (Anderson, John, Keltner, & Kring, 2001), social cohesion (Barrick, Stewart, Neubert, & Mount, 1998) and work-performance (Oh, Wang, & Mount, 2011; Zimmerman, Triana, & Barrick, 2010). Correlations between self-reports and peer-reports by acquaintances have also been found (Allik, et al., 2010; Colvin & Funder, 1991; Funder & Colvin, 1988).

Motion analyses. Several ways exist to describe physical features of the body potentially relevant to person-perception. Subjective ratings, physical measurements or computerised processing have been used. Features might include attractiveness, waist-to-hip ratios, posture cues or joint angles. Movement cues, such as lateral body sways, have further been used in the literature on person-perception. Indeed, kinematic analyses of whole-body movement have been used to discover which visual cues drive perception of gender (Kozlowski & Cutting, 1978; Mather & Murdoch, 1994; Troje, 2002), vulnerability (Gunns, et al., 2002; Johnston, Hudson, Richardson, Gunns, & Garner, 2004) or emotion (Atkinson, Dittrich, Gemmell, & Young, 2004; Pollick, Paterson, Bruderlin, & Sanford, 2001; Roether, Omlor, Christensen, & Giese, 2009). Subjective ratings of body-motion have also been used to explain trait attribution although such ratings often involve meaningful actions, such as head shaking (Ambady & Rosenthal, 1993) or an interpretation of the movement, such as youthfulness (Montepare & Zebrowitz-McArthur, 1988).

This thesis uses both subjective ratings and computer-extracted parameters related to whole-body motion. Some other physical features potentially relevant for motion, are used only for control; e.g., BMI measurements to assess group-differences. The analyses depend on the type of stimuli used.

In half of the experiments[1] (reported in Chapters II, III and IV), our stimuli consisted of different actions (Chapter II) and emotions (Chapter III), and contained two-dimensional data: patch-light and full-light displays (see Atkinson, et al., 2004; Atkinson, Tunstall, & Dittrich, 2007). Full-light displays show the whole body of the target actor but with faces obscured; patch-light displays show only patches on the targets limbs and thus contain much less visible static information (see Figure 1, examples A and B). These stimuli were created from recordings made by a single video camera. To perform motion analyses on these stimuli we used software designed to extract motion parameters from two-dimensional video clips, obtaining a measurement of the use of personal space and general space (see Movement Analyses in Chapter II).

[1] Traditional text books claim that some sort of manipulation must take place in order for a study to be called an *experiment*. However, use of the term in the literature suggests that this is becoming less restrictive. In this thesis, the term *experiment* is used when the analyses can be considered a new, semi-independent *study*, and whenever new data is collected from observers.



A) B) C)

Figure 1. Snapshot from examples of A) Full-light, B) Patch-light and C) Point-light displays. Part 1 of the thesis used full-light and patch-light movie displays as stimuli. Part 2 used point-light movie displays.

In the other half of the experiments (reported in Chapters V, VI, VII and VIII), the stimuli used were point-light displays (Johansson, 1973; see Figure 1, example C). The stimuli were created from three-dimensional motion data of people walking. Walking is cyclic, and we modelled single gait cycles through sinusoidal fitting using output from a Principal Component Analysis, resulting in four motion parameters. To label the motion parameters we also collected subjective ratings of motion (e.g., perceived amount of limb extraction). See Chapter V for detailed description of these analyses.

The Thesis Chapters: An Overview

In line with the different display formats, the experimental chapters are by and large split into two parts; in Chapters II, III and IV we present data from seven experiments on personality trait judgements based on patch-light and full-light displays, where target actors were instructed to display an emotion or an action. In Chapters V, VI, VII and VIII, we present data from a further seven experiments, this time based on point-light displays (Johansson, 1973).

Part 1: Full-light and patch-light displays of actions and emotions. In Chapter II, we present data from three main experiments (Experiments 1.1, 1.2 and 1.3) carried out using silent video clips in which targets were instructed to display a given action. Two additional experiments were carried out to obtain measurements of perceived emotional content in the video clips. The stimuli used in Experiment 1.1, 1.2 as well as the preliminary experiment were full-light displays, in which the whole body of the target actor was shown but the face was obscured. Experiment 1.3 used patch-light displays, which were versions of the full-light displays, modified so that the amount of visible static information was greatly reduced. Six rating scales were used: approachability, extraversion, neuroticism, novelty-preference, trustworthiness and warmth; high inter-rater reliability was found for all these scales. Motion parameters partly explaining this consensus were identified. We further found that type of action displayed drove personality trait judgements. Furthermore, the results confirmed that observers may extract emotional information from video clips in which no emotion was intended and that these judgements may cue personality trait judgements.

In Chapter III, we assessed whether personality trait judgements can be altered depending on intended emotion; that is, when a target actor was asked to portray a given emotion (e.g., fear). Two experiments are presented: in Experiment 2.1, we used full-light displays and in Experiment 2.2 we used patch-light displays. As in Experiments 1.1, 1.2 and 1.3, motion parameters were extracted using computer software and these were found to partly predict personality trait judgements.

In all personality trait-rating experiments we used six rating scales. In Chapter IV we asked whether these six scales were indeed separate judgements, or whether they may be underlying factors of a smaller number of higher-order constructs. To assess this question, in Experiment 3.1, we carried out a Principal Component Analysis on all the personality rating data obtained from the main experiments in Chapters II and III. The findings provided preliminary evidence that people make two judgements when asked for explicit trait impressions on whole-body motion: these were labelled dominance and valence. However, the PCA patterns differed between emotion and action clips; neuroticism in particular did not provide a consistent pattern depending the experimental design. Findings from Experiment 4.2, where we collected attractiveness judgements of the stimuli, suggested that the inconsistent loading pattern of the scale neuroticism was due to a different interpretation of the scale depending on type of behaviour displayed. Due to the different loading patterns we continued using the six traits throughout the subsequent personality trait-rating experiments to perform a similar analysis for the point-light walker stimuli.

Part 2: Point-light displays of target walkers. In Chapter V, we present the procedure of creating point-light walker stimuli, obtained through 3D motion-capture data of target individuals. Such stimuli were used so as to further reduce static information and to allow for more elaborate motion analysis. We selected walking as the type of movement because of its emotional neutrality as well as omnipresence in daily life, including situations in which trait attribution may lead to important decisions (e.g., job interviews). We further illustrate, in Experiment 4.1, how the complex motion data were reduced to a small number of components using Principal Component Analysis. These were then described by sinusoidal models with high goodness of fit. To label these components we carried out an experiment (Experiment 4.2) to obtain subjective motion descriptors of the point-light walkers. In Experiment 4.3 we show that self-reported personality traits of the target walkers were compared to these motion parameters, but no link was found. In other words, we found no evidence that someone’s personality is reflected in how they walk.

Chapter VI presents Experiment 5.1, in which we collected personality trait ratings on the point-light stimuli. These were further, in Experiments 5.2, tied to the sinusoidal motion parameters extracted by the PCA in Chapter V. To validate the findings we carried out two further rating experiments in which we manipulated the motion data in line with the correlations highlighted in Experiment 5.2. More specifically, we varied the amplitudes of the first (Experiment 5.3.1) and second (Experiment 5.3.2) motion components. The findings from these two experiments corroborated the findings from Experiments 5.1 and 5.2; this adds further support to the claim that it is the dynamic information of the point-light stimuli that drives personality trait judgements.

Having established consensus at zero acquaintance for point-light walkers in the previous chapter, we aimed, in Chapter VII, to find underlying factors other than the motion parameters explaining this consensus. First, in Experiment 6.1, we assessed whether the self-reported personality traits of the targets themselves drove the trait ratings, thus also exploring whether a kernel of truth could be established for our stimuli. Even though no link between self-reported personality and motion parameters was found in Chapter VI, we carried out this analysis due to the possibility that other physical aspects of point-light walker stimuli may not have been detected in the PCA. No accuracy was found, nor did we find any other correlations between self-reported and perceived personality. We then explored whether trait impressions other than personality could drive the consensus, by collecting data on perceived age (Experiment 6.2.1), masculinity (Experiment 6.2.2), as well as arousal, attractiveness and health (Experiment 6.2.3). We found that emotion judgements were linked to personality trait judgements, thus corroborating the experimental findings from Chapters II and III. Attractiveness and masculinity also showed potential for predicting personality trait impressions.

Chapter VIII follows up the findings presented in Chapter IV, which suggested that the six rating scales were underlying two higher-order personality traits. To do this we carried out another PCA on all the trait ratings obtained in Experiments 5.1, 6.2.1, 6.2.2 and 6.2.3. All six personality traits as well as the five predictor variables were entered into a rating-data PCA. The findings showed partial support for the two higher-order factors proposed in Chapter IV, and suggested that the scale neuroticism may be related to judgements of masculinity.

Chapter IX presents a summary of all results as well as conclusions. We propose corollaries of our findings and show how they may be of practical importance. We acknowledge some limitations in our experiments, and offer ideas for follow-up studies.

PART 1:

FULL-LIGHT AND PATCH-LIGHT DISPLAYS OF ACTIONS AND EMOTIONS

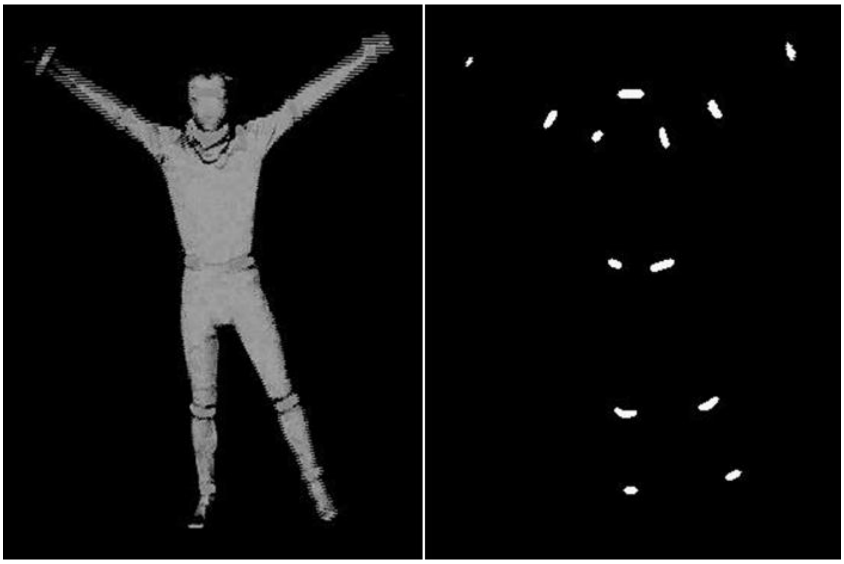
Chapter II: Judgements of Personality based on Full-Light and Patch-Light Displays of Actions

People agree with one another when making personality judgements based on pictures of faces (Bar, et al., 2006; Berry, 1990; Penton-Voak, et al., 2006; Said, Sebe, & Todorov, 2009). These stimuli are static and do not contain expressed behavioural cues, which means that people make personality judgements without context. However, there is also evidence that behavioural cues alone can lead to reliable personality trait attribution. Research has shown that simply describing the behaviour a person engages in can have an impact on observer judgements of that person, regardless of emotional expression (Uleman, et al., 2005).

People infer traits from behaviours; for instance, a single act of kindness might make an observer infer that someone is caring. In an original cued recall paradigm, Winter and Uleman (1984) showed that recall of action sentences was superior if trait words were used, rather than nouns describing the action or agent involved in the action sentence. For instance, the sentence 'the librarian carries the woman's groceries across the street' was more readily remembered with the cue *helpful* rather than *books* or *bag.* Participants were unaware of having made the trait inference. This is an example of Spontaneous Trait Inference (STI) and is evidence that people extract information beyond that available, and that we have a strong predisposition to attributing personality traits to others. A similar, if not stronger, cuing effect has been found using pictures rather than words as stimuli (Fiedler & Schenck, 2001). In that study, priming an observer with a given behaviour (e.g., a silhouette picture of someone causing harm to another) led to faster reaction times of naming an associated personality-describing adjective (e.g., 'aggressive') than for non-matching picture-trait pairs (e.g., picture of a person touching someone, with the trait adjective ‘trusting’). Using the same cued-recall paradigm, Fiedler, Schenck, Watling and Menges (2005) showed that presenting the interactions as video clips produces an equally strong STI.

The finding that interactions presented through movie clips can produce STIs led us to explore whether simpler actions, involving only one agent, could also affect explicit trait impressions. To assess this, we presented observers with silent video clips of targets engaging in different actions. In Experiments 1.1 and 1.2 we showed observers target individuals in full-light displays of body movement. Full-light displays show the whole body of the actor but the face is obscured; these video clips thus include information about body shape/form (Atkinson, et al., 2004). To make the stimuli, untrained actors were asked to engage in different actions; e.g., touching toes, and we were interested in finding out whether type of action accounted for variance in the personality ratings given by the participants. The stimuli used in Experiment 1.1 were selected through a procedure of limiting emotional content. No attempt of limiting emotional content was made in Experiment 1.2; rather, indicators of emotion were used as covariates in the analyses.

Finally, in Experiment 1.3, we repeated Experiments 1.1 and 1.2 but using patch-light versions of similar video clips (this display format resembles point-light displays; see Chapter V). During the filming, small patches were attached to the major joints and head of the actor. Viewers see only these patches move, on a black background (see Figure 2). Because the stimuli were created using a single camera, patches could be occluded by the body: e.g., if targets put their hands behind their back, the patches would not be visible to the observer. Patch-light displays contain minimal static information from the video clips but retain movement (including form-from-motion) information (Atkinson, et al., 2004; Johansson, 1973). Observers readily identify actions and emotions from such stimuli (Atkinson, et al., 2004; Atkinson, Heberlein, & Adolphs, 2007; Atkinson, Tunstall, et al., 2007). These stimuli thus allowed us to investigate whether people can make reliable personality judgements when static information about body shape/form is removed or reduced and, if so, whether the patterns of ratings were similar to those found in Experiment 1.1.



A) B)

Figure 2. Snapshot example of A) a full-light movie clip with B) corresponding patch-light version from a recording sequence of an actor performing a star jump.

There is evidence that people make personality generalisations based on the type of emotion someone is intending to display (Knutson, 1996; Montepare & Dobish, 2003; Oosterhof & Todorov, 2008). There is also evidence that people make similar generalisations based on perceived emotion even when none is intended (Said, et al., 2009) (see Chapter III). For instance, neutral face stimuli perceived to be trustworthy in a rating study were rated higher on happiness and lower on anger by separate observers (Todorov & Duchaine, 2008). People are also able to recognise expressions of emotion from dynamic displays (Atkinson, et al., 2004; Atkinson, Tunstall, et al., 2007; Dittrich, Troscianko, Lea, & Morgan, 1996; K. L. Johnson, McKay, & Pollick, 2011; Pollick, Paterson, et al., 2001). Therefore, certain measures were followed to ensure that the stimuli used in the experiments reported in this chapter contained little or no emotion. In a preliminary experiment we gathered ratings of emotional content by a separate group of observers. Stimuli that were rated towards the negative end of the scale were then selected for Experiments 1.1 and 1.3. To control for emotional content in Experiment 1.2, we collected data on arousal and valence perceived from the stimuli, again using a separate group of observers.

Consensus at zero acquaintance on non-verbal stimuli must be due to certain visible aspects of the stimuli: observer decisions must be based on something that they perceive, although observers might not be aware of using these cues. Some studies have succeeded in identifying nonverbal cues for personality trait ratings, such as symmetry in static faces: asymmetrical faces are perceived as more neurotic, less agreeable and less conscientious (Noor & Evans, 2003). Static cues for impressions of target bodies include the body mass index (BMI), which is correlated with perceived attractiveness (Swami & Tovee, 2005), as is volume to height ratio (Fan, et al., 2005; Fan, et al., 2004).

Most of the reported bodily cues for personality trait judgements are static; not many studies have described dynamic bodily cues. Attempts include a study by Montepare and Zebrowitz-McArthur (1988), who used subjective ratings of motion qualities of point-light walkers. The authors showed that the perceived ‘youthfulness’ of someone’s walk can drive impressions of the *power* of this walker. However, power does not necessarily reflect personality traits, and can also be argued to be confounded with the motion descriptions. Indeed, other research that has found consensus, and even accuracy, in trait ratings have failed to identify the actual cues that drive these judgements (Riggio, Lippa, & Salinas, 1990). An exception is Kenny et al. (1992), who found that rapid body movement was linked to impressions of extraversion, when observers were shown 20s video clips of targets who were unknowingly filmed.

In all experiments (Experiments 1.1, 1.2 and 1.3), we examined the extent to which personality ratings were influenced by two relatively low-level characteristics of the body movements (see Movement Analyses under General Methods). The action clips varied in type of movement due to the different nature of the instructions (e.g., knocking on a door requires less movement than does a star-jump) and we expected the low-level motion cues to be connected to which action was displayed. However, we also assessed whether there was any intra-actor consistency across recording conditions. Personality traits should be stable for people across different conditions (Allport, 1961) and individual differences in perceivable facial displays persists also (Ekman, 2003). The intra-individual consistency of facial expressions, this *style*, may well carry over to bodily motion. Indeed, people possess an identifiable style which allows familiar observers identify them from three-second point-light clips (Cutting & Kozlowski, 1977). We thus expected that this movement pattern, this *style*, to persist across actions to a certain extent.

It was therefore hypothesised that there would be high agreement in trait ratings between observers, in line with previous research indicating consensus at zero acquaintance. We expected high inter-rater agreement overall, but also within each action. Due to previous research on spontaneous trait inference we expected trait ratings to be related to type of action displayed. We also expected emotional content (in Experiment 1.2) to be related to trait ratings. As an exploratory element to the study, we investigated whether motion cues could be extracted to account for any of the trait ratings, irrespective of type of action displayed.

General Methods

Stimuli. Video clips in all experiments were selected from a library of stimuli (see Atkinson, et al., 2004; Atkinson, Tunstall, et al., 2007). Untrained actors, naive to the use of the recordings, were instructed to display a given action with their body. Actors wore dark grey clothes and headwear so that no part of their anatomy, including the face, was visible and they moved within a circle of approximately 2m radius. Instructions were minimal: actors were allowed to interpret the displays as they saw fit. (For a more detailed description of the recording process, see Atkinson, et al., 2004.)

Personality ratings. In all reported experiments in which we obtained personality judgements, we use rating scales based on a five-factor model of personality. The rating scales were taken from a study using similar stimuli (Heberlein, Adolphs, Tranel, & Damasio, 2004), for which the adjectives had been selected from McCrae and Costa's (1987) instrument validation study. Ostensibly, the chosen antonyms loaded strongly on their respective construct and were also easy to understand for observers whose native language may not be English.

The constructs extraversion and neuroticism were captured by the antonym pairs *shy*/*outgoing* and *calm*/*anxious*, respectively. To assess perceived agreeableness and perceived openness to experience we used the antonym pairs *unfriendly*/*friendly* and *stay-at-home*/*adventurous.* As in Heberlein et al.'s (2004) study, these scales have been labelled 'warmth' and 'novelty-preference' respectively. The construct conscientiousness, whose label has been a matter of debate (for a review, see Goldberg, 1990), was represented by the scale trustworthiness—using the antonym pair *untrustworthy/trustworthy*—due to its importance in social decision-making. Trustworthiness has also been suggested to be the rating scale that best captures a valence judgement in faces (Oosterhof & Todorov, 2009).

The present study also included a sixth scale wherein participants were asked to rate the clips on approachability, using the antonym pair *unapproachable/approachable*. This judgement frequently used in personality perception literature is of importance in social interactions and has been shown to diverge between populations; for example, patients with amygdalar lesions rate as highly approachable facial stimuli that are normally judged by healthy controls to be highly unapproachable (Adolphs, Tranel, & Damasio, 1998) and children with Williams Syndrome are more sensitive to emotional displays in their approachability ratings of pictures of faces (Frigerio, et al., 2006). Adolphs et al.'s (1998) study also found similar patterns between approachability and trustworthiness judgements, which highlights a strong link between these scales. Trustworthiness judgements capture the perceived general valence of a face, which, in turn, triggers approach/avoidance behaviour (Chen & Bargh, 1999; Oosterhof & Todorov, 2009). Therefore, whilst we expect high correlations between these scales, the processes involved to make the judgements were sufficiently different to warrant keeping both scales.

Movement analyses. In order to investigate expressiveness in body movements and gestures, Camurri, Mazzarino, Trocca, and Volpe (2002) presented methods for extracting cues from low-level motion tracking of body movements using the software EyesWeb ([version 3.3] Cohen, MacWhinney, Flatt, & Provost, 2007, www.eyesweb.org). These cues have been shown to be successful in individuating emotions (Camurri, Lagerlof, & Volpe, 2003). In that paper, cues related to the Personal Sphere and to the General Sphere were used. According to this approach, the Personal Sphere (or 'Kinesphere') denotes the space limited to the area that can be reached without changing one's stance. The General Space, in contrast, denotes all that which is beyond this sphere. In the present study, both features were extracted using the Motion Analysis library of the EyesWeb software (Cohen, et al., 2007). Some adjustments of the EyesWeb patches were performed due to the nature of our stimuli (notably length and start frame).

Related to the Personal Space, a measurement of the amount of contraction and or expansion of the body (*Contraction/Expansion Index*; CI) was extracted from the video clips. In the patch given on the EyesWeb motion analysis library (www.eyesweb.org) this parameter is computed by first calculating the ratio between the area of the silhouette of a body and its bounding region in each frame of the video clip. That is, in each of the 72 frames of our clips, the area occupied by the actor's body is divided by the area of the rectangle defined by the outer boundaries of the actor. If an actor has limbs stretched out this would manifest itself in a lower ratio, whereas a high CI_{raw} (close to 1) would indicate that an actor's limbs were kept close to the centre of the body. This parameter is then typically used to define shorter segments of movement, indicating whether a short movement is a contraction or expansion, by comparing the ratio at the beginning and the end of the segment (Camurri, De Poli, Leman, & Volpe, 2001; Camurri, et al., 2003). In the current study, however, stimuli were of a very short duration (three seconds), starting and ending with the actor facing front: no single natural movements could be defined by a start- and end-frame within the clip. Therefore, we transformed the parameter in order to give a better estimate of the use of the Personal Space. A simple arithmetic mean across the 72 frames in each movie clip would yield a measurement of the degree to which the actor's limbs had been held close to or away from the body throughout the duration of the movie clip. However, arithmetic averages, or any other linear transformations of this ratio, are sensitive to body shape/size and would also fail to gauge the degree to which the ratio changes from frame to frame. Therefore, we used standard deviation for all these ratio data points as the final parameter (referred to as CI henceforth in this thesis). The standard deviation was used because it is more sensitive to changes of limb positioning and is therefore proposed to be a better measurement of the use of Personal Space.

Secondly, a measurement of Quantity of Motion (QoM) was extracted from the video clips. In EyesWeb, QoM is based on a Silhouette Motion Image (SMI), which is an image that carries information about the change in size and position of the actor's body in the last few frames. QoM is calculated by the number of non-zero pixels of an SMI, i.e. the area that represents movement between frames. This area is then divided by the area of the most recent silhouette thus correcting for distance to camera and individual differences in body shape/size (Camurri, et al., 2003)[2]. In the present study, the QoM values were first calculated as the area of the SMI between each frame *n* and *n*-2, divided by the area of most recent silhouette. We then used the mean of these values from frame 4 to 72 to represent QoM for each clip. It must be noted that QoM and CI are not entirely independent, in that flapping of extremities will influence both measures. For that reason, QoM is sometimes referred to as a parameter of Personal Space (e.g., Camurri et al., 2001). However, as opposed to the

[2] In the classic EyesWeb QoM patch, a background extraction is first performed, thus assuming that all subsequent motion is from the actor now entering the scene. Our stimuli started in frame 1 with the actor present and a background extraction was therefore omitted.

CI, QoM is highly sensitive to movement across the visual field of the viewer. QoM could thus be said to represent the General Space as well as the Personal Space.

Experiment 1.1: Personality Ratings on Emotionally Neutral Full-Light Action Clips.

Method. Twenty-six participants (six males and 20 females) took part in the experiment (mean age = 20.8 years, *SD* = 3.15). All participants were first-year undergraduate psychology students who took part in exchange for partial course credit. This and all subsequent experiments were approved by the ethics committee of the Psychology Department of Durham University. Written, informed consent was obtained from all participants.

For all rating studies reported in this chapter, the stimuli were presented in a random order on a 12” laptop screen using the software PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993), version X, build 51, (see http://psy.ck.sissa.it). Participant responses were collected using the top row number keys of a standard QWERTY keyboard.

Fifty 3-second video clips of four untrained actors (two males) in full-light were used. During the time of recording, the actors had been instructed to engage in a particular action, facing the camera. The stimuli were selected from a larger library of body movements (see Atkinson, Heberlein, et al., 2007; Atkinson, Tunstall, et al., 2007). The selection was made through a procedure of limiting emotional content: separate observers (*N* = 15) were asked to rate how emotional they perceived the clips to be, on a scale from 0 (neutral) to 5 (highly emotional). The stimulus-set included several different instances of the same target engaging in a given action, and we chose the 50 clips out of the library that had been rated the lowest on emotional content; all clips were judged to carry little to no emotional content, with the average rating across all clips being less than 1 (M = 0.54, SD = 0.84). Due to this procedure, an uneven number of actions were represented in the selected stimulus set. These were: dig (n = 4), hop (n = 9), kick (n = 4), knock (n = 3), push (n = 7), perform a star jump (n = 5), touch toes (n = 10) or walk on spot (n = 8). Information from clothes and facial features were not available to the participants: the stimuli comprised only greyscale silhouettes of the targets on a black background (see Figure 2).

Distance to the screen was not controlled; an estimated average of 40cm means the stimuli subtended approximately 21º of visual angle (vertical). Observers rated the stimuli on the six rating scales approachability, extraversion, neuroticism, novelty-preference, trustworthiness and warmth. Five of the scales were chosen to represent the ‘Big 5’: warmth (*unfriendly*, *friendly*); trustworthiness (*untrustworthy, trustworthy*), extraversion (*shy, outgoing*), neuroticism (*calm*, *anxious*) and novelty-preference (*stay-at-home*, *adventurous*). In addition to these, a sixth scale was included: approachability (*unapproachable, approachable*). Only the antonyms were shown to participants and scales were presented in a random order for each stimulus.

Participants were given on-screen instructions in which they were told which rating scales were to be used. Participants were asked to think of the traits as being independent from one another (e.g., someone can be approachable yet introverted at the same time) and to go with their ‘gut feeling’ when making a judgement. A practice trial consisting of rating one target on all six personality trait rating scales was used. Following this, the experimental trials began. Once a stimulus appeared, participants were prompted on-screen to rate the actor on a given personality trait, on a five-point rating scale consisting of two pervasive antonyms on opposite sides. Each movie clip was played on a repeating loop until the participant had rated the depicted person on all of the six provided personality scales. The order of the trait scales was randomised, as was the order of the presentation of the movie clips. As for all the following experiments reported in the thesis, participants were fully debriefed of the hypotheses of the study upon completion of the experiment.

Results. Inter-rater reliability tests showed agreement between observers was good for the trait scales approachability, extraversion, novelty-preference and warmth: all Cronbach's αs > .80 and the effective reliability, *R*, was > .80 on these scales as calculated by the Spearman Brown formula (Rosenthal, 1973). Reliability for the scales trustworthiness (α = .64; *R* = .65) and neuroticism (α = .66; *R* = .66) were acceptable. Examining the action types individually, inter-rater agreement was low (mean α = .57; mean *R* = .49). That is, within an action type, participants did not fully agree which clips were rated high or low on the different personality traits.

A multivariate analysis of covariance was carried out with the different action clips as the random factor (i.e., an item analysis). Since there was an unequal number of actions per actor, actor ID was included as a fixed factor. The MANCOVA thus included actor ID and type of action as a fixed factor, with Quantity of Motion (QoM) and Contraction/Expansion Index (CI) as covariates. Results revealed that this model accounted for a significant proportion of the variance (Wilks' Lambda: $F(6,15) = 91.01$, $p < .001$; $\eta_p^2 = .97$). The test showed that, having controlled for both CI and QoM, type of action had a significant impact on the participants' personality trait judgements, $F(42,73.8) = 2.91$, $p < .001$, $\eta_p^2 = .54$. Examining the rating scales individually, type of action was found to have a significant impact on all personality scales, all Fs(7,20) > 3.10, all ps < .05, all η_p^2s > .52; highest effect sizes were found for extraversion ($\eta_p^2 = .80$) and novelty-preference ($\eta_p^2 = .68$). Figure 2 shows a summary of all personality ratings across type of action displayed by the target in the clips. As can be seen from this figure, star jumps were rated the highest on all rating scales apart from neuroticism. No other strong patterns can be detected, although it does appear that an action that was rated low on one scale was also rated low on other scales.

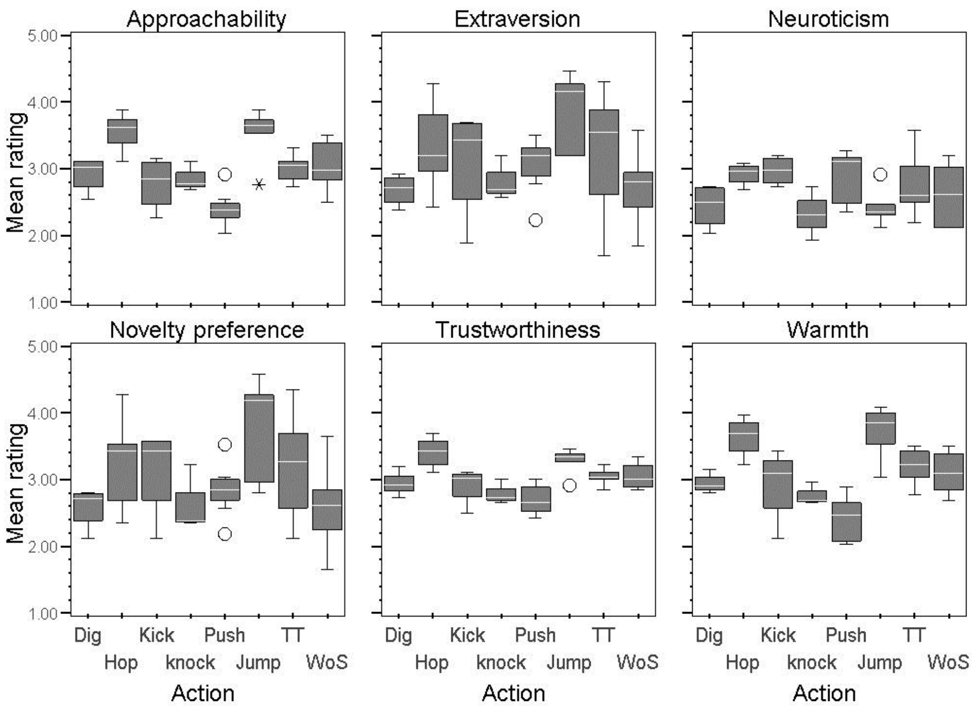


Figure 3. Mean ratings for all personality scales and across actions for emotionally neutral full-light action clips. TT = Touch Toes. WoS = Walk on Spot. Jump = Star jump.

The results also indicate that QoM accounted for a significant proportion of the variance in the personality ratings, $F(6,15) = 22.18$, $p < .001$; $\eta_p^2 = .90$. Examining the personality scales individually, QoM was found to have a significant impact on personality ratings on all scales apart from neuroticism ($p = .72$). The effect size was greatest for ratings of extraversion, $F(1,20) = 136.93$, $p < .001$; $\eta_p^2 = .87$, and novelty-preference, $F(1,20) = 79.32$, $p < .001$; $\eta_p^2 = .80$. The remaining personality traits saw much lower effect sizes (all η_p^2s $< .36$). Significant correlation coefficients were all positive; that is, the more movement was seen within a clip, the higher this clip was rated on all five personality scales. It is difficult to assess whether this overall positive direction was due to different actions varying in QoM *and* in trait ratings, as sample sizes within each action were small. However, the signs of the coefficients were mostly positive; i.e. consistent with the overall direction (see Table 1).

Table 1

Correlations (ρ) between Quantity of Motion and Personality Ratings for Emotionally Neutral Full-Light Action Clips

	Intended Action								Across
Rating scale	Dig	Hop	Kick	Knock	Push	Jump	TT	WoS	Actions
Approach.	.11	.50	.60	−.50	.04	.50	−.20	.31	.47**
Extraversion	−.40	.82**	.80	.50	.71	.87	.81**	.76*	.75**
Novelty-pref.	−.20	.82**	.95	.50	.71	.99**	.92**	.95**	.74**
Trustworth.	.32	.68*	.80	−.50	−.07	.10	.33	.26	.52**
Warmth	.40	.62	.60	.50	−.02	.80	.43	.20	.60**

Note. Jump = star jump, TT = touch toes, WoS = walk on spot. * $p < .05$; ** $p < .01$. No correction of multiple comparisons is performed, because the effects are already identified using multivariate statistics.

No significant main effect was found for CI on trait ratings, $F(6,15) = 0.55$, $p = .763$, $\eta_p^2 = .18$. Furthermore, no main effect was found for Actor ID, although this was close to reaching significance, $F(18,42.9) = 1.65$, $p = .09$, $\eta_p^2 = .39$.

Discussion. Type of action was found to have a significant, albeit weak, impact on personality ratings. This held, above all, for novelty-preference and extraversion judgements. These two scales also saw the highest inter-rater agreement through the different actions.

The weaker impact of type of action on personality ratings may be due to the uneven number of clips within each action category. To rectify this, we replicated Experiment 1.1 using an even number of stimuli for each action. As in Experiment 1.1, the clips had no *intended* emotion, but, since no effort was made to reduce this, they nonetheless varied in emotional content. This thus allowed us to assess whether emotional cues can also be used as a basis for personality trait judgements in video clips where no emotion was intended.

Experiment 1.2: Personality Ratings on Full-Light Action Clips with Emotional Content

Experiment 1.1 showed that type of action has an impact on personality ratings for stimuli with little or no emotional content. In Experiment 1.2 arousal and valence ratings were gathered, and used as 1) covariates for the analyses of how type of action influences personality ratings and 2) whether observers also use emotional cues to fuel personality ratings when these emotional displays are not intentional.

Method. A total of 20 new participants (five males) took part in the experiment (mean age = 19.5 years, *SD* = 1.05). As before, participants were sampled from psychology undergraduate students who took part in exchange for partial course credit.

This experiment used 72 video clips similar to those used in Experiment 1.1 and selected from the same stimulus library (Atkinson et al., 2007). The 72 clips consisted of 12 clips from each of six actions (hop, knock, push, star jump, touch toes and walk on spot). Six actors (three males and three females) engaged in each of the actions twice. In half the instances, the actor was asked to perform the action with low intensity, the other with high intensity. This was done to ensure a range of motion within the stimuli. These selected clips were shown to a separate group of observers (*N* = 16, mean age = 21.4 years, *SD* = 8.8) who made judgements of arousal and valence on a 5-point rating scale (mean arousal: 3.15, *SD* = 1.02, mean valence: 3.14, *SD* = 0.81). All further aspects of the stimuli were identical to those used in Experiment 1.1. Due to a limited pool of stimuli, 30 of the clips in Experiment 1.2 had been used Experiment 1.1.

The design was identical to that of Experiment 1.1 with the following exception: The main independent variable was the type of action displayed by the actor during the encoding. There were six levels of this variable: hop, knock, push, star jump, touch toes and walk on spot. Two control variables were included: ratings of arousal and valence, which were treated as 1) covariates in the analyses of effect of action on personality ratings and 2) predictors of personality ratings in the analyses of effect of emotion on personality ratings. As before, two low-level motion parameters, Quantity of Motion (QoM) and Contraction/Expansion Index (CI), were entered as 1) covariates for the analyses of the effect of emotion, and 2) as intrinsically predictive variables. The procedure was identical to that in Experiment 1.1.

Results and Discussion. There was high agreement between observers for ratings of the video clips (all *R*s > .75; all *α*s > .79). There was generally high agreement within each action (mean *α* = .68, mean *R* = .70).

A multivariate analysis of covariance was carried out, including type of action and actor ID as fixed factors, and QoM, CI, arousal and valence as covariates. This model was found to account for a significant proportion of the variance of the six personality scales, $F(6,27) = 21.20$, $p < .001$, $\eta_p^2 = .83$. Type of Action had a significant impact on personality judgements, $F(30,110) = 4.12$, $p < .001$, $\eta_p^2 = .45$. This held for all traits apart from Neuroticism. The traits Approachability, $F(5,32) = 10.54$, $p < .001$, $\eta_p^2 = .62$, and Warmth, $F(5,32) = 7.99$, $p < .001$, $\eta_p^2 = .56$, were the scales that saw the most variance explained by type of action, whereas Trustworthiness, $F(5,32) = 4.27$, $p < .01$, $\eta_p^2 = .40$, saw the least variance explained by type of action. The identity of the actor in the movie clips did not have a significant impact on personality ratings, Wilks' lambda: $F(30,110) = 1.47$, $p = .08$, $\eta_p^2 = .24$.

Of the motion parameters, QoM was the only independent variable found to influence trait judgements, $F(6,27) = 3.21$, $p < .05$, $\eta_p^2 = .42$. This was significant for ratings on the scale extraversion only, $F(1,32) = 5.63$, $p < .05$, $\eta_p^2 = .15$. As can be seen in Table 2, the more movement was seen in a clip, the higher this clip was rated on extraversion, and this held for all actions. Note that for this and subsequent analyses, only traits that were significantly affected by the motion parameters are shown. CI was, once again, found to have no main effect on personality ratings, $F(6,27) = 0.97$, $p = .47$, $\eta_p^2 = .18$.

Table 2

Correlations (ρ) between Quantity of Motion and Extraversion Ratings for Full-Light Action Clips with Emotional Content

	Type of action						Across
Rating scale	Hop	Knock	Push	SJ	TT	WoS	Actions
Extraversion	$.70^{*}$	$.77^{**}$	$.77^{**}$	$.93^{**}$	$.83^{**}$	$.92^{**}$	$.72^{**}$

Note. SJ = star jump, TT = touch toes, WoS = walk on spot. * $p < .05$; ** $p < .01$.

The covariates arousal, $F(6,27) = 10.77$, $p < .001$, $\eta_p^2 = .71$, and valence, $F(6,27) = 7.53$, $p < .001$, $\eta_p^2 = .63$, were both found to account for a substantial amount of variance in the personality rating data. In other words it would appear that there is some link between perception of emotional cues and judgements of personality traits from whole-body movements that were made without the intention to portray emotional states. The effect sizes were considerably higher for these emotion cues than they were for type of action, actor ID or movement parameters. Arousal had a significant impact on all scales (all *p*s < .015), with the strongest effect sizes found for novelty-preference ($\eta_p^2 = .51$), neuroticism ($\eta_p^2 = .48$) and extraversion ($\eta_p^2 = .42$). Linear correlations showed that, for the scales extraversion, neuroticism and novelty-preference, the relation was positive; that is, the more *aroused* a clip had been judged, the more *outgoing*, *anxious* and *adventurous* the target on that clip was rated. This held across all conditions, i.e. regardless of type of action. For the other scales, the direction was ambiguous; for instance, although there was a positive correlation between ratings on warmth and arousal in which actors were asked to ‘walk on the spot’ ($\rho = .77$, $p < .001$), the correlation was negative for the clips in which the actor was instructed to ‘knock’ ($\rho = -.81$, $p < .001$; see Table 3).

Table 3

Correlations (ρ) between Arousal and Personality Ratings for Full-Light Action Clips with Emotional Content

	Type of action						*Across*
Rating scale	Hop	Knock	Push	Jump	TT	WoS	*Actions*
Approachability	.10	−.80**	−.79**	.09	−.33	.55	*.20*
Extraversion	.91**	.91**	.74**	.90**	.85**	.77**	*.86***
Neuroticism	.19	.90**	.87**	.85**	.90**	.17	*.71***
Novelty-preference	.89**	.87**	.78**	.89**	.92**	.70*	*.86***
Trustworthiness	.30	−.86**	−.24	−.01	.12	.38	*.15*
Warmth	.52	−.81**	−.52	.59*	.40	.77**	*.38***

Note. Jump = star jump, TT = touch toes, WoS = walk on spot. * $p < .05$; ** $p < .01$.

Valence accounted for a significant proportion of the variance in ratings on the scales warmth, $F(1,32) = 34.70$, $p < .001$, $\eta_p^2 = .52$, approachability, $F(1,32) = 24.40$, $p < .001$, $\eta_p^2 = .43$, trustworthiness, $F(1,32) = 16.24$, $p < .001$, $\eta_p^2 = .34$, and neuroticism, $F(1,32) = 11.53$, $p < .01$, $\eta_p^2 = .27$. The higher a clip was rated for valence, the higher this clip was rated on all three personality scales. This held within all actions: that is, no significant correlation was negative (see Table 4).

Table 4

Correlations (ρ) between Valence and Personality Ratings for Full-Light Action Clips with Emotional Content

	Type of action						Across
Trait scale	Hop	Knock	Push	Jump	TT	WoS	Actions
Approachability	.10	−.28	.13	.00	−.40	.58*	.57**
Neuroticism	.48	−.09	−.19	.77**	.93**	.37	.28*
Trustworthiness	.33	.16	.01	.00	.08	.47	.49**
Warmth	.37	.33	.39	.68*	.41	.76**	.75**

Note. Jump = star jump, TT = touch toes, WoS = walk on spot. * $p < .05$; ** $p < .01$

QoM correlated with arousal (ρ = .76, $p < .001$) and valence ratings (ρ = .66, $p < .001$). Due to this and the strong correlation between emotion ratings and personality ratings, a partial correlation between QoM and extraversion was carried out (controlling for arousal and valence). Since the multivariate test controls for inter-scale covariance, we know that the relation between QoM and personality ratings is not driven by emotion ratings. However, the coefficients were strongly reduced, suggesting that the relation was to some degree affected by emotion ratings (see Table 5).

Table 5

Partial Correlations (ρ) between QoM and Extraversion, controlling for Arousal and Valence, for Full-Light Action Clips with Emotional Content

	Type of action						Across
Trait scale	Hop	Knock	Push	Jump	TT	WoS	Actions
Extraversion	.01	.82**	.63	.59	.14	.44	.30*

Note. Jump = star jump, TT = touch toes, WoS = walk on spot. * $p < .05$; ** $p < .01$.

Figure 4 shows that, as was found for Experiment 1.1, star-jumps were rated highest on all scales apart from neuroticism. This fits nicely with literature on perception of aesthetical movements from dance stimuli; it has been found that vertical displacements (jumps) are more aesthetically pleasing than movements with little vertical displacement (Calvo-Merino, Jola, Glaser, & Haggard, 2008). The same trend of cross-scale convergence was again found (i.e. clips that were rated low on one scale were typically rated low on other scales). The exception was for the scale neuroticism; no patterns similar (or inverse) to other personality traits were identified.

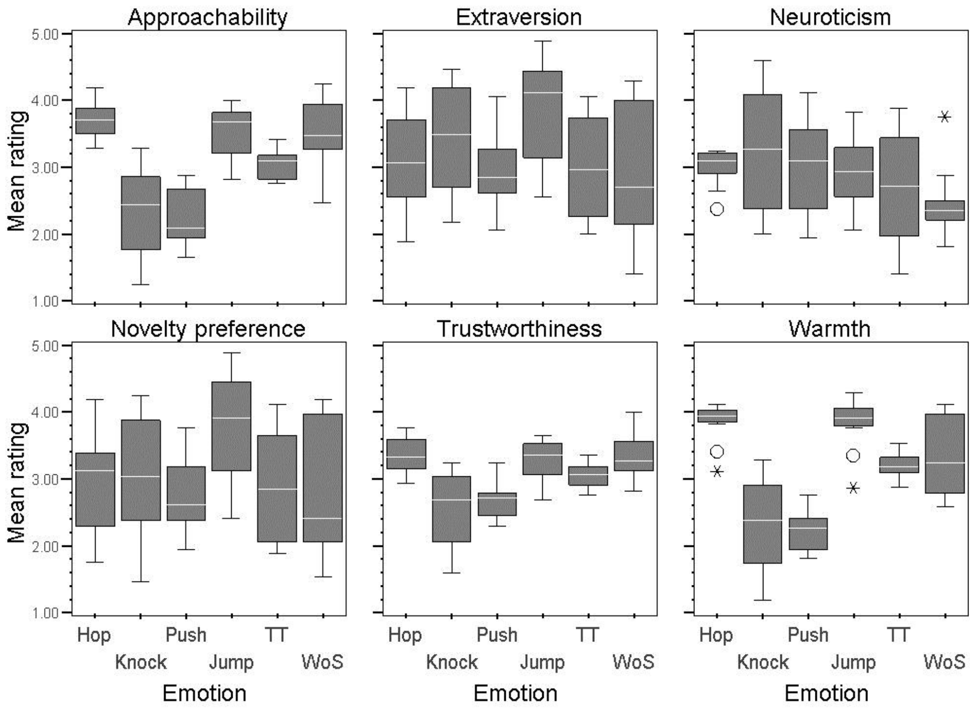


Figure 4. Mean ratings for all personality scales and across actions for full-light action clips with emotional content. TT = Touch toes, WoS = Walk on spot, Jump = Star-jump.

Discussion: Full-Light Experiments

Experiments 1.1 and 1.2 have shown that people tend to agree in their judgements of personality traits when they are presented with impoverished stimuli of full-light video clips. The data further suggest that—consciously or otherwise—judgements are partly based on which action the target of a video clip portrays; that is, if an actor displays a given action, this affects how others judge this actor's personality. Experiment 1.1 showed that the type of action that a target engages in can affect personality judgements of that target when emotional content is reduced. Findings from Experiment 1.2 showed that emotional cues can be used to fuel trait impressions also when the target individual does not intentionally express an emotion. In other words, people do not rely on the perception of emotional expression-related cues alone when they make personality judgements, but may—perhaps unconsciously—use other first-order cues in the type of movement in which the to-be-judged person is engaged.

Findings from both experiments indicate that the movement parameter QoM accounted for variance in the trait ratings given by the participants. Although QoM varied between conditions, the analyses showed that the use of this cue transcended type of action. It is important to note that the relation between QoM and personality ratings was reduced when the ratings of arousal and valence were partialled out. In other words it could be that, if we indeed have found a low-level motion cue that can predict personality ratings, the relation might not be entirely direct. Rather, the observer might use the movement cue to assess the emotional state of an actor and then base the personality trait judgements on this decision.

The findings thus far indicate that observers agree when attributing personality traits to people based on full-light recordings of their body movements lasting only three seconds. One low-level feature drawn from motion analysis of the stimuli, QoM, was shown to explain some variance in the judgements. Personality traits were also shown to be affected by type of action. Whether these judgements are based on movement alone is questionable, since full-light displays still show considerable static information (e.g., body shape). A third experiment was therefore run to assess whether similar findings would be obtained for patch-light displays, in which the available static information is substantially reduced.

Experiment 1.3: Personality Ratings on Patch-Light Action Clips

Method. Twenty-three new participants (four males) took part in the experiment (mean age = 20.0 years, *SD* = 3.0). As in previous studies, participants were all undergraduate psychology students who took part in the study in exchange for partial course credit.

A total of 56 patch-light movie clips were used, comprising the movements of four untrained actors (two males); all stimuli were drawn from the set developed by and reported in Atkinson et al (2007). Thirty-eight of the clips were based on the video clips used in Experiments 1.1 and 1.2. The clips were all three seconds long. During the filming, actors had been asked to engage in one of the following eight actions: dig (n = 6), hop (n = 7), kick (n = 7), knock (n = 8), push (n = 8), star jump (n = 7), touch toes (n = 5) or walk on spot (n = 8). Full-light versions of the stimuli were rated by a separate group of observers (N = 15; see Experiment 1.1), who judged the clips on emotional content. On a scale from 0 (neutral) to 5 (highly emotional), the mean emotionality rating of the full-light versions of the selected clips was 0.76 (SD = 0.49); in other words, the stimuli carried little or no emotional content. Each actor performed each action at least once, and no actors performed any actions more than twice. Any further apparatus was identical to the previous experiments reported in this chapter (see Experiment 1.1).

The design was identical to that of Experiment 1.1. Due to availability the following eight actions were used: dig, hop, kick, knock, push, star jump, touch toes or walk on spot. As before, the low-level motion parameters, QoM and CI were entered as 1) covariates for the analyses of the effect of emotion, and 2) as intrinsically predictive variables. The values that had been extracted from the full-light versions of the stimuli were used (for a detailed description see Movement Analyses under General Methods). The procedure was identical to that used in the previous experiments reported in this chapter (see Experiment 1.1 for a full description).

Results and Discussion. There was acceptable reliability in personality judgements of the video clips. The lowest agreement, across type of action, was found for Neuroticism ($\alpha = .691$; $R = .676$); the highest, for Extraversion ($\alpha = .930$; $R = .926$). Reliability was lower within actions (mean $\alpha = .64$; mean $R = .63$), although coefficients were high for the actions kick ($\alpha = .70$; $R = .76$) and knock ($\alpha = .84$; $R = .81$). This indicates that the overall α was strong not so much due to agreement about which *individual clip* was approachable but, rather, which *action* was approachable; e.g., that actors who were hopping were more approachable than when they were pushing (see Figure 5).

A multivariate analysis with actor ID and type of action as fixed factors, and with QoM and CI as covariates, accounted for a significant proportion of the variance, $F(6,17) = 74.14$, $p < .001$, $\eta_p^2 = .96$. Type of action had a significant impact on the personality ratings, $F(42,83.19) = 2.49$, $p < .001$, $\eta_p^2 = .47$, and this held for all scales apart from neuroticism: all $F(3,22) > 3.3$, all ps $< .05$, all $\eta_p^2 > .51$. The strongest effect sizes were seen for warmth ($\eta_p^2 = .71$) and novelty-preference ($\eta_p^2 = .62$). Star-jumps were rated higher than were other actions on all scales apart from neuroticism. These findings are thus partially in line with findings from Experiments 1.1 and 1.2.

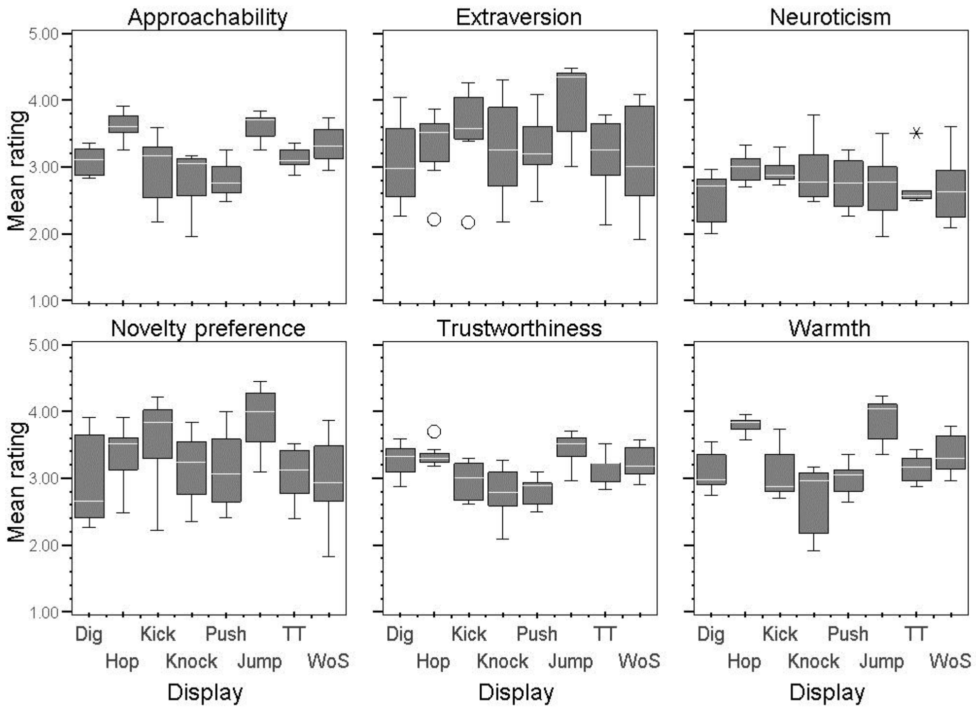


Figure 5. Mean ratings for all personality scales and across actions for patch-light action clips. TT = touch toes, WoS = Walk on spot, Jump = Star jump.

Personality judgements partially depended on which actor figured in the clip, $F(18,48.6) = 3.32$, $p < .001$, $\eta_p^2 = .52$. This held only for three of the traits: neuroticism, $F(3,22) = 4.81$, $p < .05$, $\eta_p^2 = .40$; extraversion, $F(3,22) = 8.11$, $p < .001$, $\eta_p^2 = .52$; and novelty-preference, $F(3,22) = 8.73$, $p < .001$, $\eta_p^2 = .54$.

Both motion parameters explained a significant proportion of the variance. Quantity of Motion, $F(6,17) = 6.78$, $p < .001$, $\eta_p^2 = .71$, had an impact on the scales extraversion, $F(1,22) = 22.54$, $p < .001$, $\eta_p^2 = .51$, and novelty-preference, $F(1,32) = 38.69$, $p < .001$, $\eta_p^2 = .64$. The more movement there was in the general space of an actor, the higher the ratings of these two scales. None of the actions saw a correlation in the opposite direction (see Table 6).

Table 6

Correlations (ρ) between QoM and Personality Ratings for Patch-Light Action Clips

	Intended Action								Across
Trait scale	Dig	Hop	Kick	Knock	Push	Jump	TT	WoS	Emotions
Extraversion	.49	.68	.21	.31	.72*	.82*	.99**	.32*	.57**
NP	.49	.68	.21	.50	.86**	.89**	.99**	.35	.67**

Note. NP = novelty-preference, Jump = star jump, TT = Touch Toes, WoS = Walk on Spot. * $p < .05$; ** $p < .01$.

CI, which had no significant impact on personality ratings in Experiments 1.1 or 1.2, did account for some variance in the current experiment, $F(6,17) = 2.82$, $p < .05$, $\eta_p^2 = .50$. This held for all scales (all Fs > 7.8, all ps < .01) but effect sizes were small (all η_p^2s < .35). The only significant correlations between CI and trait ratings were found for extraversion ($\rho = .52$, $p < .001$) and novelty-preference ($\rho = .52$, $p < .001$). In other words, the more use of the personal space of a target, the more outgoing and adventurous that target was judged to be. The direction of this relation varied within actions (see Table 7) but few actions saw a consistent negative correlation (with exception, perhaps, of *hop*).

Table 7

Correlations (ρ) between CI and Personality Ratings for Patch-Light Action Clips

	Intended Action								Across
Trait scale	Dig	Hop	Kick	Knock	Push	Jump	TT	WoS	Actions
Approach.	−.06	.14	.26	−.88**	.06	.41	.67	−.30	.05
Extrav.	.06	−.50	.74	.62	.67	.14	.56	.55	.52**
Neurotic.	.09	−.71	.37	.63	.35	.00	.10	−.06	.11
NP	.06	−.50	.74	.49	.54	.21	.56	.36	.52**
Trust.	−.34	.04	−.15	−.69	.36	.36	−.13	−.33	−.04
Warmth	.22	.41	.30	−.65	.18	.57	−.05	−.17	.13

Note: Approach. = approachability, Extrav. = extraversion, Neurotic. = neuroticism, NP = novelty-preference, Trust. = trustworthiness, Jump = star jump, TT = Touch Toes, WoS = Walk on Spot. * $p < .05$; ** $p < .01$.

Correlational analyses showed that there was a link between how a clip was rated in its full-light version and how it was rated in its patch-light version. All ρs > .43, all *p*s < .01, with the highest coefficients found for novelty-preference (ρ = .82) and extraversion (ρ = .87**)**.

The findings from Experiment 1.3 show good reliability for personality ratings on patch-light movies in which targets engaged in one of a fixed number of actions. Type of action was a strong predictor of trait ratings of the video clips, also when other motion cues, and personal style, were controlled for. This adds to the literature showing that dynamic displays of action can affect personality judgements (Fiedler, et al., 2005), confirming that this also holds for patch-light displays. Indeed, it has been shown that people are accurate in identifying actions from patch-light stimuli (Atkinson, 2009; Atkinson, Tunstall, et al., 2007; Runeson & Frykholm, 1983).

There was moderate to strong agreement in trait ratings also *within* actions. In other words, participants were also using a cue other than type of action to make trait judgements. The multivariate tests identified a number of valid cues, including two motion parameters. One of these, QoM, quantifies movement within the general space; the other, CI, quantifies movement within personal space.

Having corrected for all these variables, the identity of the actor also affected personality ratings. This may suggest that people have a personal 'movement style', also beyond the motion parameters CI and QoM. This style might then cue personality judgements by observers.

General Discussion

The above experiments have shown that people readily make personality trait judgements based on nonverbal cues from the body only when seeing a target for only three seconds displaying an action, and that people converge in their judgements. Experiments 1.1 and 1.2 showed that this is true for full-light displays in which the targets are asked to display a given action. Experiment 1.3 expanded on these findings, yielding similar results also for patch-light displays of similar movie clips. This suggests that observers need not rely on cues from body shape/size, clothes or voice when judging someone's personality traits. There were also significant correlations between personality ratings given for patch-light versions and full-light versions of the same video recordings.

The cues beyond type of action were stronger predictors of personality ratings for patch-light stimuli than they were for full-light stimuli. For instance, the second low level motion characteristic, Contraction/Expansion Index (CI), had a significant impact on personality ratings for patch-light displays only (Experiment 1.3). In other words, for patch-light versions of clips in which actors displayed a given action, the use of personal space was able to predict how observers judged the targets. Intuitively, as there are fewer static form cues in the patch-light stimuli, participants were forced to rely more on dynamic characteristics to cue personality judgements.

We have shown that low-level features extracted by means of motion analysis on simple digital video footage can predict personality judgements. Performing further motion tracking analyses is likely to identify other physical characteristics that cue trait judgements. In turn, these findings could be of interest to animators and makers of computer games. If these extracted cues are reliable indicators of personality traits, it should be possible to manipulate these cues to produce avatars and other computer models whose movements convey particular traits. Indeed, systematically manipulating body motion cues using computer-based motion capture and animation is the logical next step for assessing the contribution of those cues to trait judgements. Bruderlin and Williams (1995) presented an approach with the potential of making such manipulations, with the most promising being motion displacement mapping of different joints (e.g., moving of wrists to a higher impact point during a knock-on-the-door). Studies have also successfully changed perception of emotions by manipulating body movements (Amaya, Bruderlin, & Calvert, 1996; McDonnell, Jörg, McHugh, Newell, & O'Sullivan, 2009; Roether, et al., 2009; Roether, Omlor, & Giese, 2008). For instance, Amaya et al (1996) showed that emotional ‘transforms’ could be successfully applied to a neutral motion, resulting in movements closely resembling the original emotional displays. To calculate such transforms, the authors used the difference between an action (e.g., kicking) performed without any expressed emotion and the same action performed by the same actor expressing an emotion (e.g., anger). These findings add to the potential suitability of computer animation techniques to change perceived personality from human avatars by altering body motion cues of these avatars.

A further assessment of how personality judgements can be made on the basis of dynamic cues should be made by limiting the number of conditions, for instance by using stimuli of only one type of action. This could further allow for situations in which observer judgements are more pertinent. For instance, although it might be helpful for the lead role candidate during an audition, it is not realistic for most job candidates to engage in actions or overtly portray an emotion during an interview. Future studies could thus include credible movements that occur within interactions; for example, how people walk into an interview room and take a seat, or how they raise their glass during a first date.

This study suggests that people reliably draw personality trait inferences from body movement. Several cues have been identified to influence trait judgements, notably type of action or emotion intended by the actor and quantity of movement seen by the observer. Judgements made on patch-light versions correlated with those made on full-light versions of the same video recordings. This suggests that body movement may influence personality trait judgements also when some static information is accessible to the observer. The findings may be used to manipulate perceived personality traits in computer modelled avatars or during human interactions, although further research must be carried out to explore this hypothesis.

The measurements of emotion—as captured through ratings of arousal and valence—had an impact on trait ratings in study 1.2. This is in line with previously cited literature using photographs of faces, showing that observers base personality trait judgements on perceived emotional content even when no emotion is intended (Said, et al., 2009). However, it is interesting that this held also for impoverished stimuli depicting whole-body movements (no faces) with distinct types of action. The link between perceived personality and displayed emotion is further assessed in Chapter III.

Chapter III: Judgements of Personality based on Full-Light and Patch-Light Displays of Emotions

Emotion is expressed through several channels, including the face, body and voice, and there is considerable evidence that people can accurately identify emotional displays based on individual channels alone (for reviews see e.g., Atkinson & Adolphs, 2005; Bänziger, Grandjean, & Scherer, 2009; Ekman, 2003). Most nonverbal research reporting emotion recognition has used photographs of faces as stimuli. However, there is support for the claim that people are able to recognise at least the basic emotions based on dynamic (e.g., Atkinson, et al., 2004; Atkinson, Tunstall, et al., 2007; Dittrich, et al., 1996) and static (e.g., Coulson, 2004; Schindler, Van Gool, & de Gelder, 2008) whole-body stimuli absent of facial cues, and even based on dynamic cues from arm movements alone (K. L. Johnson, et al., 2011; Pollick, Paterson, et al., 2001). This skill seems to require little practice; children as young as four years show an ability to detect emotion from body movement alone (Boone & Cunningham, 1998).

Personality is considered a trait stable across time (Costa & McCrae, 1988; Loehlin & Martin, 2001). This contrasts with the typical view of emotions: someone who is angry in one moment can be happy in another. However, since one of the main definitions of personality is the ‘dynamic organization of ... psychophysical systems that determine [a person’s] characteristic behavior and thought’ (Allport, 1961, p. 28), it is easy to see how emotion and personality are linked. The lexical approach to classifying personality (see Chapter I) also confirms this connection; for instance, the statement ‘I rarely feel fearful’ has a negative loading on the construct Neuroticism in the NEO Five Factor Inventory (Costa & McCrae, 1992).

Driven by the observation that people make inferences from faces pertaining to important social decisions (e.g., Little, et al., 2007), and by the limited accuracy of such inferences (Todorov, et al., 2005), a number of recent studies have highlighted the link between emotional expression and personality judgements in faces. Most notably, this research has shown that perceived emotion can affect personality trait judgements. Oosterhof and Todorov (2008) showed that perceived emotional displays can be changed by varying trustworthiness features in faces: untrustworthy faces appeared angry and trustworthy faces appear happy. Similarly, Montepare and Dobish (2003) found that type of emotional facial display predicted ratings of dominance and affiliation. In that study, happy faces were deemed to have the most desirable personality traits, whereas angry expressions were rated high in dominance and low in affiliation. These correlations persisted even when the stimuli were created without any intended emotion displayed, i.e. when the faces were neutral. Facial cues that resemble emotion expressions have been found to be associated with perceived trustworthiness: V-shaped eyebrows and more ∩-shaped mouths have been found to yield lower trustworthiness ratings (Todorov, Baron, & Oosterhof, 2008). This is suggestive that trustworthiness judgements are based on cues that resemble expressions signalling whether a person can be approached or should be avoided (Oosterhof & Todorov, 2009; Todorov, et al., 2008).

There is also evidence that perceived personality traits can affect emotion intensity judgements; for instance, manipulating trustworthiness has been found to change perceived anger (in angry faces) and happiness (in happy faces) (Oosterhof & Todorov, 2009). It seems, then, that the perceptual link between emotion and personality is bidirectional. This raises the question as to whether emotion and trait impressions are not processed independently but, rather, that they rely on the same processes. However, although trait inferences might be partly confounded with emotion perception, studies of abnormal populations do suggest a dissociation between personality and emotion judgements (e.g., Porter, Coltheart, & Langdon, 2007) and there is evidence that ratings of personality and of emotion from point-light displays of body movements may rely on separate, though inter-connected, cortical regions (Chouchourelou, Matsuka, Harber, & Shiffrar, 2006; Heberlein, et al., 2004; Heberlein & Saxe, 2005). This suggests that, while people might use emotion cues to judge personality traits, trait inferences may also rely on other cues.

We conducted two experiments to build on findings from Chapter II, by assessing whether personality ratings are influenced by the intended type of emotion displayed in 1) full-light displays and 2) patch-light displays of body movement. As in Chapter II, we identified motion parameters that could predict such trait ratings. Qualitatively and quantitatively different movements manifest the type and intensity of emotions expressed with the body (Atkinson, et al., 2004; Brownlow, Dixon, Egbert, & Radcliffe, 1997; Camurri, et al., 2003; Wallbott, 1998). For example, Camurri et al. (2003) found that fear and grief were typically expressed with little use of the personal space compared to other emotions, ‘limbs often close to the centre of gravity.’ (p. 223) Anger and joy, on the other hand, yielded greater measurements of general space compared to other emotions. However, there are inter-individual differences in expressiveness of emotion (Allport, 1961), and links between facial expression and personality have been shown to exist (Kring & Sloan, 2007); for instance, intensity and duration of positive facial expressions are positively correlated with extraversion and negatively correlated with neuroticism. Furthermore, it is known that people are able to recognise themselves and known peers from their movements (Prasad & Shiffrar, 2009). This suggests that individuals carry a personal behavioural style and we thus had reason to believe that the motion parameters identified in Chapter II could also predict personality trait ratings with emotional stimuli.

Experiment 2.1: Personality Ratings on Full-Light Emotion Clips.

Method. Twenty-two participants (five male and 17 female; mean age = 18.8 years, *SD* = 0.7) took part in the study. All participants were undergraduate psychology students who took part in exchange for partial course credit. Informed consent was obtained from all participants.

A total of 50 three-second movie clips in full-light were used, each clip depicting one of 10 different actors (five males and five females). During the time of encoding, the actors were instructed to display a given emotion: anger, disgust, fear, happiness or sadness (see Figure 6). These clips were selected from a library of stimuli presented by Atkinson et al. (2004) and Atkinson et al. (2007). The same selection procedure as that adopted in Atkinson, Tunstall, et al. (2007) was employed. That is, selection was done on the basis that the intended emotions were well recognised; ten versions were included for each emotion, but as a result of the selection procedure the 10 actors were not necessarily represented an equal number of times in each emotion category.

A) B)

Figure 6. Snapshot example of A) a full-light movie clip with B) corresponding patch-light version, from a recording sequence of an actor displaying the emotion happiness.

All other aspects of the Method were identical to the experiments reported in Chapter II.

Results. Inter-rater reliability tests showed that participants had a high level of agreement in how they rated the video clips on all six scales: effective reliability (R) was > .82 on all scales as calculated by the Spearman Brown formula (Rosenthal, 1973), all αs > 0.82. Reliability was moderate to high also *within* emotions: the mean effective reliability for all six personality scales was calculated for each of the emotions. There was high effective agreement overall (R = .71; α = .68), with the lowest agreement seen for clips intended to portray disgust (R = .47; α = .46) and the highest for clips intended to portray happiness (R = .84; α = .85)

A multivariate analysis of covariance was carried out, with type of emotion and actor ID as fixed factors, and QoM and CI as covariates. The MANCOVA accounted for a significant proportion of the variance: Wilks' Lambda: $F(6,38) = 409.33$, $p < .001$; $\eta_p^2 = .99$. The test further revealed that, having controlled for the differences of both CI and QoM within the emotions displayed, and the different number of individual actors within each condition, the type of emotion had a significant impact on the trait judgements, $F(24,25.6) = 15.15$, $p < .001$, $\eta_p^2 = .91$. Emotion had a significant impact on all personality scales (all *p*s < .001), with strong effect sizes (all η_p^2s > .87). Happy stimuli were rated higher than other emotions on all personality scales apart from neuroticism. Anger clips were rated the lowest on approachability, trustworthiness and warmth (see Figure 7).

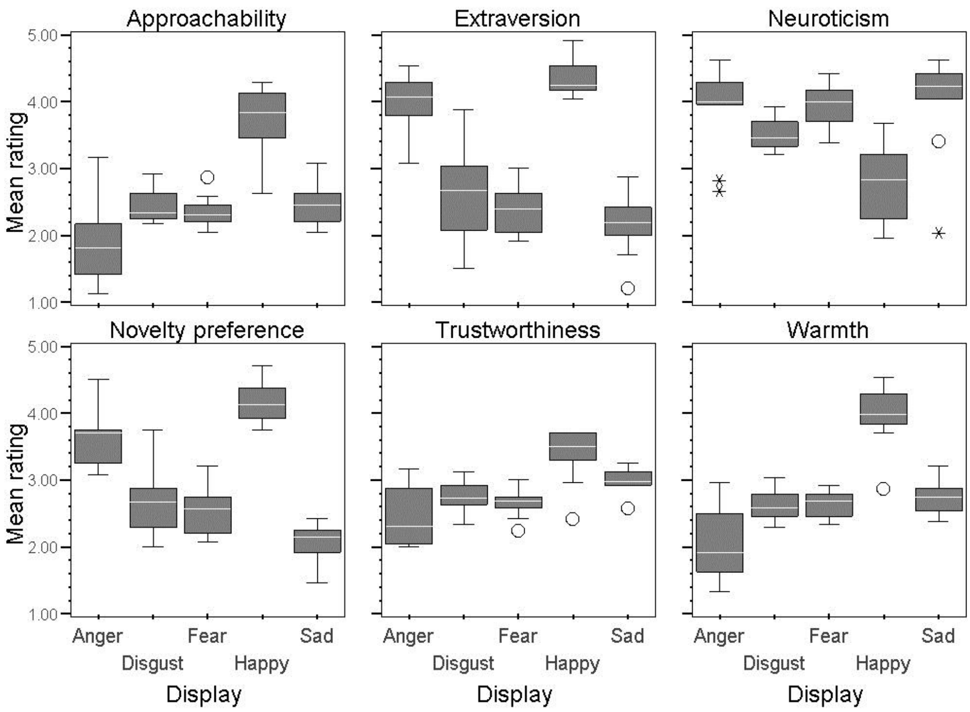


Figure 7. Mean ratings for all personality scales and across emotions for full-light stimuli.

QoM too accounted for a significant proportion of the variance in the ratings given by the participants, and the effect size was strong, $F(6,7) = 5.18$, $p < .05$; $\eta_p^2 = .82$. The influence of QoM on the trait judgements was significant for the scales neuroticism, $F(1,12) = 10.73$, $p < .01$; $\eta_p^2 = .47$, and novelty-preference, $F(1,12) = 6.31$, $p < .05$; $\eta_p^2 = .35$, only. The higher was QoM the higher were the judgements on both scales. In other words, the more a target moved the more anxious and adventurous this target was rated. To address the possibility of inflated correlations, coefficients *within* each emotion were calculated. Although these varied highly for the different emotions, none of the emotions saw a correlation in the opposite direction (see Table 8).

Table 8

Correlations (ρ) between Quantity of Motion and Personality Ratings for Full-Light Emotion Clips

	Intended emotion					Across
Trait Scale	Anger	Disgust	Fear	Happy	Sad	Emotions
Neuroticism	.31	.23	.38	.15	$.74^{**}$	.02
Novelty-preference	.40	$.93^{***}$	.41	.37	.01	$.61^{***}$

Note. ** $p < .01$; *** $p < .001$

Which actor figured on the video clip had a significant impact on the personality ratings, $F(54,40.3) = 2.52$, $p < .01$, $\eta_p^2 = .72$. Examining the scales individually, this was significant for approachability, $F(9,12) = 3.08$, $p < .05$, $\eta_p^2 = .70$, and neuroticism, $F(9,12) = 4.88$, $p < .01$, $\eta_p^2 = .78$, only. Rerunning a MANCOVA revealed that this effect was not due to sex of the actor, $F(6,33) = 1.24$, $p = .31$.

Discussion. The results from Experiment 2.1 indicate that there is agreement between viewers as to the personality traits of the stimulus targets, even after very short exposure to these clips. The findings also show that the intended type of emotion has an impact on personality ratings. It is plausible that the identified inter-rater agreement of personality traits (across all clips) is carried by an agreement of what aspects of movement convey *emotional* information. This is in line with previous research; indeed, the stimuli used had been chosen for their high emotion recognition rate (Atkinson, et al., 2004) and studies show a tendency for overgeneralisation of trait impressions based on emotion perceptions (Knutson, 1996; Montepare & Dobish, 2003). It thus appears that viewers pick up on emotional cues and make judgements (explicitly or implicitly) of personality traits based on these emotion judgements; in other words, the link between personality judgements and physical aspects of the stimuli is mediated by emotion.

However, the results also suggest that participants used cues other than the intended emotion when making personality judgements. Firstly, there were high inter-rater reliability measures within each individual emotion. That is, not only did participants agree that, say, ‘happy clips’ were more approachable and less anxious than, say, ‘sad clips’; they also agreed on which of the happy clips were more approachable or calmer than other happy clips. Secondly, a main effect on personality ratings was found for the motion parameter Quantity of Motion, having corrected for emotional display. This is further evidence that cues other than type of emotion are at play in personality judgements from emotional body movements, and that these cues can be extracted by means of computer-based motion analyses (as we will explore further in Chapter V.) This is in line with evidence that people rely on distinct mechanisms for judging personality and emotion (Heberlein, et al., 2004; Heberlein & Saxe, 2005).

Experiment 2.2: Personality Ratings on Patch-Light Emotion Clips

This experiment was a replication of Experiment 2.1; stimuli were obtained from the same set (Atkinson, Tunstall, et al., 2007) in which untrained actors, wearing white patches on their joints, had been asked to display a certain emotion. As was the case for Experiment 1.2 (Chapter II), movie clips had been altered to make the white patches the only visible cues; observers would thus only see white patches move around on a black screen (see *Figure 6*).

Method. Twenty-four participants (five males and 19 females) took part in the experiment (mean age = 19.7 years, *SD* = 1.5). All participants were psychology undergraduate students who took part in exchange for partial course credit. Written, informed consent was obtained from all participants.

Fifty video clips were used. The video clips were patch-light versions of the same footage as those used in Experiment 2.1 (see Figure 1); thus, expressions of the emotions anger, disgust, fear, happiness and sadness were used. Further required apparatus was identical to that used in Experiment 2.1.

The design and procedure were identical to those of Experiment 2.1. As was the case for Experiment 1.2, motion parameters (CI and QoM) were extracted from the full-light version of the clips.

Results and Discussion. Inter-rater reliability was high across all stimuli (all Rs > .87; all αs > .87). Cross trait-scale agreement was moderate (mean R = .70; mean α = .71) but varied between emotions: agreement was high for happiness (R = .85; α = .86) and anger (R = .91; α = .90) stimuli but low for other emotions (Rs < .63; αs < .68).

A MANCOVA was carried out to test whether type of emotion affected personality ratings. Actor ID was also entered as a fixed factor and CI and QoM were entered as covariates. The model accounted for a significant proportion of the variance, $F(6,7) = 282.95$, $p < .001$, $\eta_p^2 = .99$. In line with the findings from the full-light video clips, type of emotion had a significant impact on the personality ratings, $F(24,25.6) = 8.94$, $p < .001$, $\eta_p^2 = .86$. This held for all personality scales (all *p*s < .001, all η_p^2s > .87) with particularly strong effect sizes for warmth, $F(4,12) = 105.33$, $p < .001$, $\eta_p^2 = .97$, and approachability, $F(4,12) = 50.13$, $p < .001$, $\eta_p^2 = .94$. As in Experiment 1, happiness stimuli were rated higher than stimuli expressing other emotions on all personality scales apart from neuroticism. Similarly, anger stimuli were rated the lowest on approachability, trustworthiness and warmth (see Figure 8). The pattern of the rating data was similar to that found in Experiment 2.1; that is, the emotions that scored high on an individual scale for full-light scored high also for patch-light displays (cf. Figure 7).

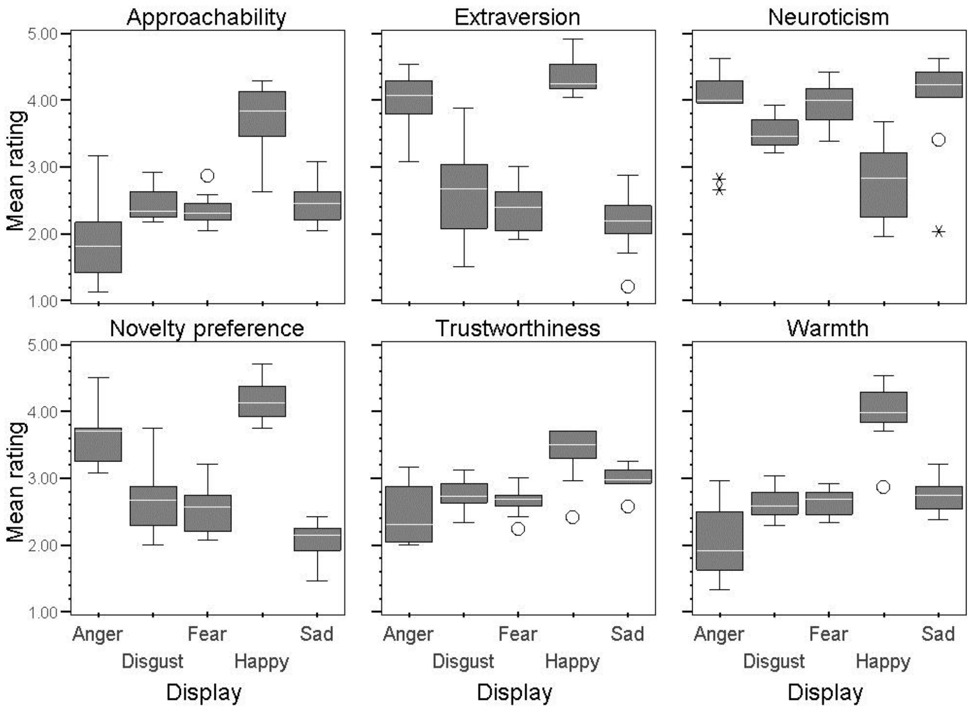


Figure 8. Mean ratings for all personality scales and across emotions for patch-light stimuli.

The motion parameter QoM explained a significant proportion of the variance in the personality ratings, $F(6,7) = 4.67$, $p < .05$, $\eta_p^2 = .80$. This held for four of the scales: approachability, $F(1,12) = 6.86$, $p < .05$, $\eta_p^2 = .36$, extraversion, $F(1,12) = 6.33$, $p < .05$, $\eta_p^2 = .35$, neuroticism, $F(1,12) = 14.07$, $p < .01$, $\eta_p^2 = .54$, and novelty-preference, $F(1,12) = 11.93$, $p < .01$, $\eta_p^2 = .50$. The effect sizes were largest for the two scales neuroticism and extraversion, which is in line with the findings obtained for the full-light versions of the stimuli in Experiment 2.1. Correlation coefficients indicate that the order of the relationship was positive for the scales extraversion, neuroticism and novelty-preference. That is, the more general movement was seen, the more extraverted, anxious and adventurous were the judgements. However, for the scale approachability the order was reversed; target individuals who moved more were in fact rated as *less* approachable (see Table 9).

Table 9

Correlations (ρ) between Quantity of Motion and Personality Ratings for Full-light Emotion Clips

	Intended emotion					Across
Trait Scale	Anger	Disg.	Fear	Happy	Sad	Emotions
Approachability	−.42	−.17	.13	−.12	$-.72^{*}$	<.01
Extraversion	$.64^{*}$	$.82^{**}$	.29	−.13	.53	.64
Neuroticism	$.68^{*}$	.16	.49	.31	$.71^{*}$	.08
Novelty-preference	.46	$.72^{*}$	.02	.13	.49	.59

Note. Disg. = disgust. * $p < .05$; ** $p < .01$.

As was the case in Experiment 2.1, the second motion parameter, CI, was found to have no impact on personality ratings, $F(6,7) = 0.28$, $p = .93$. Which actor figured in the video clip had a significant impact on personality ratings, $F(54,40.29) = 2.17$, $p < .01$, $\eta_p^2 = .69$. As was the case in Experiment 2.1, this impact of actor ID was significant for the scales approachability, $F(9,12) = 3.21$, $p < .05$, $\eta_p^2 = .71$, and neuroticism, $F(9,12) = 9.24$, $p < .001$, $\eta_p^2 = .87$. Additionally, in Experiment 2.2, actor ID influenced ratings for the scale trustworthiness, $F(9,12) = 4.14$, $p < .05$, $\eta_p^2 = .79$.

The findings described thus far show that personality ratings for patch-light displays follow a similar pattern as for full-light displays, when targets are instructed to portray a given emotion. This was further supported by a correlation analysis: coefficients across type of emotion were all positive, significant and moderate to strong (see Table 10). As Table 10 shows, there was also some relationship between personality ratings also *within* emotions, especially for anger, disgust and happiness clips. That is, if a clip in which an actor displayed the emotion anger was rated as, say, unapproachable for the full-light version, this clip was rated as unapproachable in the patch-light version too ($\rho = .93$, $p < .01$)

Table 10

Correlations (ρ) between Personality Ratings given for Full-Light and Patch-Light Emotion Clips

	Intended emotion					Across
Trait Scale	Anger	Disg.	Fear	Happy	Sad	Emotions
Approachability	.83**	.75*	.62	.62	.53	.77**
Extraversion	.74*	.86**	.58	.68*	.50	.93**
Neuroticism	.64*	.51	.81**	.93**	.76*	.86**
Novelty-preference	.93**	.88**	.36	.76*	.55	.94**
Trustworthiness	.59	.69*	.02	.73*	.14	.70**
Warmth	.85**	.75*	.30	.74*	.69*	.83**

Note. Disg. = disgust. * $p < .05$; ** $p < .01$.

The effect of emotion on personality ratings was thus similar for patch-light as for full-light clips. This may come as little surprise, given that observers generally are accurate in identifying intended emotions in dynamic patch-light stimuli (Atkinson, et al., 2004; Atkinson, Tunstall, et al., 2007; Dittrich, et al., 1996). However, this study is (to our knowledge) the first to show that overgeneralisations occur also when participants are exposed to dynamic emotion displays in which static information from body shape/size is substantially reduced.

There was inter-rater reliability also *within* emotions. This agreement suggests that participants used other cues to fuel trait ratings, which provides some support for reliable movement cues that are unrelated to emotion. Multivariate analyses did indeed find a main effect of motion characteristics that persisted once type of emotion had been corrected for. The use of ‘general space’ (Camurri, et al., 2003) had a significant impact on two of the scales in the full-light stimuli (neuroticism and extraversion) and these also saw the largest effect sizes for the patch-light stimuli.

The findings from Experiment 2.2 follow the same pattern as in Experiment 2.1. In other words, the different emotions were rated consistently regardless of whether the clips were presented in full-light or patch-light format.

General Discussion

The findings of the two experiments reported in this chapter show that, without cues from voice, facial expressions, and visual manifestations (hair style, clothing, jewellery etc.), people are able to rely on body motion to judge personality traits when targets display emotions. The inter-rater agreement, found within emotions and across both types of display format, adds to a large body of research indicating consensus at zero acquaintance (Albright, et al., 1997; Ambady & Rosenthal, 1993; Kenny, et al., 1992; Penton-Voak, et al., 2006).

Type of emotion had an impact on personality trait ratings for both display formats. This is in line with previous research showing that observers can extract emotional information from dynamic displays (Atkinson, et al., 2004; Atkinson, Tunstall, et al., 2007; Camurri, et al., 2003; Dittrich, et al., 1996; Pollick, Paterson, et al., 2001), and confirms that people may indeed use impressions of emotion to fuel personality attribution (Knutson, 1996; Montepare & Dobish, 2003; Oosterhof & Todorov, 2009; Said, et al., 2009).

Different emotions manifest different physical motion features (Atkinson, et al., 2004; Camurri, et al., 2003; Wallbott, 1998) and so we expected the motion parameters to be affected by emotion display type. Interestingly, QoM had an impact on trait ratings also *within* the emotion types. This was the case for the scales neuroticism and novelty-preference for both display formats; in addition to approachability and extraversion in patch-light stimuli only. Contrary to what was found in Experiments 1.1, 1.2 and 1.3 (see Chapter II), QoM was *negatively* correlated with approachability. QoM was also positively correlated with neuroticism. This may be taken as preliminary evidence that, whilst more movement leads to higher ratings on socially desirable traits when the target engages in an action, it may lead to *lower* ratings of such traits when the target engages in an emotion. This was most evident for the emotions anger and sadness, and (for full-light clips only) fear, indicating that more movement when engaging in negative emotions leads to ‘negative’ trait attribution.

The confirmation that emotion affects trait attribution raises the question of whether such inferences are valid; that is, whether they have some basis in the real world. Intuitively, a person who experiences happiness when causing pain to others is not likely to possess the same personality traits as someone who experiences sadness in that same situation, so it is clear that for personality trait attribution to be accurate the observer must take context into account. However, there also seems to be some support for a general tendency to express emotions, irrespective of context, being related to personality. Personality literature using self-reports suggests that a non-contextual tendency to express happiness is associated with neuroticism (e.g., Costa & McCrae, 1992), whilst a tendency to express anger is negatively associated with agreeableness (e.g., Ashton & Lee, 2009). It has also been shown that someone who successfully expresses anger is likely to be dominant (Friedman, DiMatteo, & Taranta, 1980). The finding that observers readily use emotional expressions to cue personality trait judgements when no context is provided does therefore not automatically suggest that such inferences are incorrect.

The between-rater agreement found within emotions, as well as the impact of the motion parameters irrespective of display, support individual differences in expression of emotion. Research on emotional expressiveness has shown a link between personality and the ability to express different emotions (Friedman, et al., 1980; Riggio & Friedman, 1986). Expressiveness has also been found to affect personality trait impressions; males displaying more postural shifts and head movements were judged as more extraverted, as were females who were judged high in facial expressiveness by trained observers (Riggio & Friedman, 1986). It could be that observers are able to pick up on cues that lead to implicit impressions; personality traits seem to affect physiological states of actors when they are asked to portray given emotions: extraverted individuals who were asked to express anger, but not happiness, saw increased heart rate compared to introverted individuals (Bono & Vey, 2007). This could explain the obtained inter-observer agreement found within the emotions.

Several questions have not been answered in these chapters. The inter-rater agreement as well as effects of motion parameters and display instructions on six personality traits is all very well, but we do not know whether such decisions occur in real-life. Plenty of evidence points towards implicit and explicit trait judgements interacting (for a review, see e.g., Uleman, et al., 2005) and impressions gained from faces have been found to impact on social choices (Little, et al., 2007; Mattes, et al., 2009; Todorov, et al., 2005; van't Wout & Sanfey, 2008). Body-motion alone has been found to affect social decisions such as voting choice (Kramer, et al., 2010). This supports the claim that trait impressions are very real. However, to our knowledge, no study has successfully tied such social decisions to all constructs of the Big Five, and it is uncertain whether people indeed make as many as five, let alone six, judgements when observing others. To answer this question, we performed, in Chapter IV, a principal component analysis of the trait ratings obtained across all experiments in Chapters II and III to find whether higher-order constructs exist in trait impressions based on whole-body movement.

Consensus among our observers leads one to ponder whether there is accuracy in the trait impressions. This has not been assessed in the first two experimental chapters; no trait ratings were obtained from the actors. Although inter-actor differences were shown in both trait impressions and body motion cues, we believe that assessment of accuracy should be performed on stimuli of targets engaging in neutral, naturally-occurring movements, because the instruction itself (i.e. display given actions or emotions) affects trait impressions. To further assess these questions we created new stimuli, as reported in Chapter V. Here, we created emotionally neutral stimuli of targets walking. A new display format was used, namely, point-light (Johansson, 1973), in order to further limit static information. Furthermore, these new stimuli were created using 3D motion-capture technology, in order to obtain richer motion data allowing for more complete investigations of motion cues.

Chapter IV: Pattern of Personality Ratings on Full-Light and Patch-Light Action- and Emotion Displays

We have shown that observers make reliable trait judgements when faced with limited access to bodily motion, for movie clips in which a target individual displays an action (Chapter II) or an emotion (Chapter III). The experiments were designed such that observers made judgements on six rating scales, and they were given unlimited time to make these judgements.

But how many judgements do people make, implicitly or explicitly, when forming an impression of others? The number of traits required to describe someone's personality is a matter of debate (see Digman, 1997, for a review) and the number of judgements made by an observer is no more straight-forward. It has been found that explicit trait judgements correlate with each other (e.g., Oosterhof & Todorov, 2008). Such degrees of colinearity is thought to occur because of a generalisation of traits; specifically, an unconscious bias of one impression, such as attractiveness, fuelling other impressions, often referred to as the Halo effect (Dion, Berscheid, & Walster, 1972).

To assess whether colinearity also occurred for our stimuli correlation coefficients were calculated using the rating data obtained in Chapters II and III. Four stimulus sets were used: full-light actions, patch-light actions, full-light emotions and patch-light emotions and a correlation matrix were calculated for each of these sets. Table 11 shows the mean coefficients across the four data sets; these data indicate that there is a high degree of colinearity between approachability, trustworthiness and warmth (mean ρ = .85) as well as high correlation between extraversion and novelty-preference (ρ = .96).

Table 11

Mean Correlation Coefficients (ρ) between Personality Ratings, across Five Experiments

Rating Scale	Appr.	Extr.	Neur.	NP	Trust.
Extraversion	.11				
Neuroticism	−.48*	.36*†			
Novelty-preference	.15	.96***	.39*†		
Trustworthiness	.82***	.00	−.32*	.00	
Warmth	.91***	.17	−.33	.21	.81***

Note. Coefficients between ratings given on the personality traits was first calculated for four separate display categories (emotion or action) X (full-light or patch-light). Numbers in the table represent the mean of these four coefficients. Appr. = approachability, Extr. = extraversion, Neur. = neuroticism, NP = novelty-preference, Trust. = trustworthiness. * average $p < .05$. *** average $p < .001$. † absolute coefficients are used.

This colinearity leads to the question as to whether observers effectively make six trait judgements. An alternative view is that the personality trait impressions are based on underlying latent variables: a decision on one dimension that might then cue the other impressions. When high degrees of colinearity between dependent variables exist, it is common practice to investigate whether some of these should be combined, or even dropped from the data analysis.

One way of reducing dimensionality of data sets is by use of Principal Component Analysis. This is a mathematical tool used in exploratory analyses where the goal is to reduce a large number of possibly correlated variables into fewer orthogonal variables, or components. The components are ordered such that the first one accounts for the most amount of variance in the data.

PCA has been used successfully for trait scales obtained for facial stimuli. Oosterhof and Todorov (2008) used principal component analysis to reduce fifteen variables down to only two components, together accounting for 81.6% of the variance. Since all positive judgements (e.g., sociable) had a positive loading and all negative judgements (e.g., unhappy) had a negative loading on the first component, this was labelled 'valence'. Dominance, confidence and aggression ratings had positive loadings on the second component and this was therefore labelled 'dominance'. The two labels make intuitive sense because an important first decision when you see someone is whether to approach or avoid this person based on perceived intention (valence) and ability (dominance) to cause harm (Fiske, Cuddy, & Glick, 2007; Oosterhof & Todorov, 2008; Todorov, et al., 2008). Two dimensions of personality are reported in other forms of personality research, such as asking people to use a list of adjectives to describe people (Rosenberg, Nelson, & Vivekananthan, 1968). Factor analysis always involves subjective interpretation and as a consequence labels chosen may differ from study to study. The second component is often referred to as *competence*; however it can still be considered a decision regarding someone's ability to cause harm (Fiske, et al., 2007).

To address whether the same pattern is found for trait impressions made on whole-body movement, we conducted a PCA (Experiment 3.1) on the data presented in Chapters II and III. Further analyses were carried out because a discrepancy was found between the action and emotion stimuli in that the scale neuroticism did not show the same pattern of loading on the components. In addition to motion analyses, which assessed whether the movement aspects are judged differently depending on the type of behaviour, we collected attractiveness judgements (Experiment 3.2). This was done to gauge whether there was a different interpretation of the scale neuroticism depending on whether the displayed behaviour was an action or an emotion.

Experiment 3.1: Principal Components Analysis on Personality Ratings on Full-Light and Patch-Light Stimuli

Method. Data from the five experiments presented in Chapters II and III were collated. Experiments 1.1 and 1.2 had 30 movie clips in common and, due to a strong between-experiment correlation on how these clips were rated on the respective trait scales (see Table 12), the arithmetic mean of the rating scores between the two experiments was employed for the analyses.

Table 12

Correlations (ρ) between Personality Ratings on Full-light Action Clips Presented in Experiments 1.1 and 1.2

Rating Scale	ρ
Approachability	.85
Extraversion	.84
Neuroticism	.71
Novelty-preference	.86
Trustworthiness	.79
Warmth	.89

Note. Correlations are based on 30 full-light stimuli that were presented in both studies. All correlations significant at the .001 level.

The stimuli were grouped into four categories for the purpose of the PCA: two display formats (full-light or patch-light) × two behaviour types (emotion or action). Due to the way in which the stimuli were collected, the categories were of unequal sample size: 148 action clips (92 full-light and 56 patch-light) and 100 emotion clips (50 full-light and 50 patch-light) were used.

In order to reduce the dimensionality of the data we used Principal Component Analysis (PCA). This procedure identifies clusters of variables, or components, that account for large amounts of variance in the data. Mathematically, the components are the eigenvectors of the covariance matrix; the components are ordered so that the first has the largest possible variance, i.e. accounting for the most amount of the variability. Components are by definition orthogonal to each other. The data were normalised and standardised at the beginning of the PCA procedures. Varimax rotation was used to maximise the variance explained by the components.

Five PCAs were carried out. The means of the six rating scales across all observers in Experiment 1.1 (n = 26), Experiment 1.2 (n = 20), Experiment 1.3 (n = 23), Experiment 2.1 (n = 22) and Experiment 2.2 (n = 24) were entered into a PCA, thus consisting of the whole dataset for all 248 clips (hereafter referred to as an *overall PCA*). In addition, the four categories were treated separately to assess whether the same pattern of results was obtained across the types of display format or instruction. Notably, we were interested to see whether there were differences in judgement patterns for action versus emotion clips, due to the data from the experiments presented in Chapters II and III suggesting that motion parameters affect valence ratings differently depending on whether a target actor displayed an action or an emotion.

Results and Discussion.

Overall PCA. The Overall PCA, which was carried out on the trait ratings for all movie clips from Experiments 1.1, 1.2, 1.3, 2.1 and 2.2 (N = 248), yielded a two-factor solution. The first component accounted for 55.9%, and the first two components together accounted for 87.6% of the variance in the rating data.

A factor loading solution using Varimax rotation showed that each of the rating scales had a strong loading on one, and only one, of the components (see Table 13). PC1 saw a strong positive loading of approachability, emotional stability (neuroticism reversed), trustworthiness and warmth whereas PC2 saw a strong positive loading of extraversion and novelty-preference. PC1 and PC2 could thus be said to fit with Oosterhof and Todorov's (2008) valence and dominance components, respectively.

Table 13

Factor Loadings of the Rating Scales on all Stimuli across Display– and Behaviour Type for PCA on all Six Rating Scales

	PC	
Rating scale	1	2
Approachability	**.98**	.15
Extraversion	.09	**.98**
Neuroticism	−**.74**	−.30
Novelty-preference	.13	**.98**
Trustworthiness	**.92**	−.10
Warmth	**.95**	.14

Note. PC = principal component. Boldface indicates factor loading stronger than .40.

The factor scores from PCAs are calculated through complex statistics; in our case multiplying eigenvalues of the covariance matrix with the coefficients of the principal components. This procedure will lead to unequal trait scale loadings depending on the dataset. To assess whether the factor scores from the PCA could be approximated by linear combinations of the trait scales used in the rating experiments we created new variables (*constructs*) by calculating the arithmetic means of the scales associated with the two components. A *valence construct* was calculated as the mean of approachability, emotional stability (the complement of neuroticism, i.e. 6-x, where x is the neuroticism rating, henceforth referred to as 'complemented'), trustworthiness and warmth. A *dominance construct* was obtained by calculating the mean of ratings given on extraversion and novelty-preference. These two variables were strongly correlated with their respective factor score only (see Table 14). A reliability measure of the individual scales in each respective component was calculated. This revealed very high inter-item reliability for both the valence construct ($\alpha = .90$) and the dominance construct ($\alpha = .98$). Furthermore, the constructs were only weakly correlated with each other ($\rho = .22$, $p < .001$).

Table 14

Correlations (ρ) between Component Factor Scores and 1) Valence and 2) Dominance Constructs

	Construct	
PC	Valence	Dominance
1	$.97^{***}$	<.01
2	.03	$.99^{***}$

Note. *** $p < .001$.

PCAs on action clips. Due to the findings that trait ratings vary by type of action (see Chapter II) and emotion (see Chapter III), the two behaviour types were treated separately. The PCAs run on action clips indicated that the six rating scales (approachability, extraversion, neuroticism, novelty-preference, trustworthiness and warmth) could be reduced to only two components. Across full-light and patch-light displays, the first two PCs accounted for 90.25% of the variance (see Table 15).

Table 15

Variance Explained by Components for PCA on all Six Rating Scales for Action Clips

	Full-light		Patch-light		Mean	
PC	Perc. var.	Cum. var.	Perc. var.	Cum. var.	Perc. var.	Cum. var.
1	51.0%	51.0%	47.1%	47.1%	49.1%	49.1%
2	39.1%	90.1%	35.6%	87.7%	37.4%	86.5%

Note. Only principal components with eigenvalue > 1 are shown. PC = principal component number, Perc. var. = percentage of variance explained, Cum. var = cumulative percentage of variance explained.

However, the datasets of the action clips were complex: Neuroticism had an absolute loading of above .40 on more than one of the components; as can be seen in Table 16, neuroticism had a negative loading on PC1 and a positive loading on PC2.

Table 16

Factor Loadings of Six Rating Scales According to Display Format for Action Clips

	PC1			PC2		
Rating scale	FL	PL	*M*	FL	PL	*M*
Approachability	**.98**	**.95**	**.97**	−.02	−.03	−.03
Extraversion	.07	.09	.08	**.97**	**.95**	**.96**
Neuroticism	**−.47**	−.40	**−.43**	**.69**	**.55**	**.62**
NP	.12	.13	.12	**.97**	**.96**	**.96**
Trustworthiness	**.94**	**.93**	**.93**	−.11	−.06	−.09
Warmth	**.96**	**.93**	**.95**	.12	.16	.14

Note. PC = principal component, NP = novelty-preference, FL = full-light displays, PL = patch-light displays, *M* = mean loading across display type. Boldface indicates loadings stronger than .40.

Because the neuroticism ratings did not have a clear loading on only one of the components, new PCAs were carried out for the action clips, this time excluding neuroticism ratings. This analysis again showed a two-component solution; PC1 and PC2 together now accounted for 94.6% of the variance in the remaining data (see Table 17).

Table 17

Variance Explained by Components for PCA on all Rating Scales excluding Neuroticism for Action Clips

	Full-light		Patch-light		Mean	
PC	Perc. var.	Cum. var.	Perc. var.	Cum. var.	Perc. var.	Cum. var.
1	57.0%	57.0%	55.2%	55.2%	56.1%	56.1%
2	39.3%	96.3%	37.6%	92.8%	38.5%	94.6%

Note. Only principal components with eigenvalue > 1 are shown. PC = principal component number, Perc. var. = percentage of variance explained, Cum. var = cumulative percentage of variance explained.

The patterns of the loadings of the five rating scales were consistent with those found for the initial PCAs, indicating a valence construct consisting of approachability, trustworthiness and warmth ratings, and a dominance construct consisting of extraversion and novelty-preference ratings. The pattern was similar across display format; that is, whether full-light or patch-light stimuli were used did not appear to have an impact on the judgement (see Table 18).

Table 18

Factor Loadings of Five Rating Scales According to Display Format for Action Clips

	PC1			PC2		
Rating scale	FL	PL	*M*	FL	PL	*M*
Approachability	**.98**	**.96**	**.97**	.03	.01	.02
Extraversion	.01	.03	.02	**.99**	**.98**	**.99**
NP	.06	.07	.07	**.99**	**.98**	**.99**
Trustworthiness	**.96**	**.94**	**.95**	−.07	−.03	−.05
Warmth	**.97**	**.94**	**.96**	.15	.19	.17

Note. PC = principal component, NP = novelty-preference, FL = full-light displays, PL = patch-light displays, *M* = mean loading across display type. Boldface indicates loadings stronger than .40.

A reliability analysis of the scales underlying the construct showed good reliability for both valence (full-light $\alpha = .95$; patch-light $\alpha = .93$) and dominance (full-light $\alpha = .99$; patch-light $\alpha = .97$), confirming that the rating scales indeed were variables underlying their respective constructs.

As was done for the Overall PCA, we constructed new variables through linear combinations of the rating data in order to see whether the component scores could be approximated using the raw rating data. Dominance was defined as the mean of extraversion and novelty-preference, whereas valence$_A$ was defined as the mean of approachability, trustworthiness and warmth. These constructs were strongly correlated with the factor scores that they were meant to approximate; that is, the valence$_A$ construct correlated with PC1 factor scores whilst the dominance construct correlated with the PC2 factor scores. Neuroticism ratings correlated with the second principal component, meaning that perceived anxiousness to a degree overlapped with perceived dominance for action clips. However, the coefficients were not strong, and we therefore have grounds to treat neuroticism as a separate judgement from the dominance and valence constructs. See Table 19. The two constructs only weakly correlated with each other (ρ = .22, $p < .01$).

Table 19

Correlations (ρ) between Component Factor Scores and A) the Valence Construct, B) the Dominance Construct and C) Neuroticism Ratings

	A) Valence			B) Dominance			C) Neuroticism		
PC	FL	PL	*M*	FL	PL	*M*	FL	PL	*M*
1	.99***	.99***	.99***	.17	.11	.15	-.22*	-.20	-.21
2	.19	.13	.16	.99***	.99***	.99***	.53***	.38**	.46***

Note. PC = principal component number, FL = full-light displays, PL = patch-light displays, *M* = Mean coefficient across display category. * $p < .01$; ** $p < .01$; *** $p < .001$.

PCAs on emotion clips. Two PCAs were carried out on the trait ratings for the emotion clips (again treating full-light and patch-light samples separately). These both yielded a two-component solution. On average, the first two components explained 90.3% of the variance of the data (see Table 20). Once again, the pattern was similar across display format.

Table 20

Variance Explained by Components for PCA on all Six Rating Scales for Emotion Clips

	Full-light		Patch-light		Mean	
PC	Perc. var.	Cum. var.	Perc. var.	Cum. var.	Perc. var.	Cum. var.
1	60.0%	60.0%	58.7%	58.7%	59.4%	59.4%
2	31.6%	81.6%	30.2%	88.9%	30.9%	90.3%

Note. Only principal components with eigenvalue > 1 are shown. PC = principal component number, Perc. var. = percentage of variance explained, Cum. var = cumulative percentage of variance explained.

Using Varimax rotation, the PCA showed that each of the six rating scales had a strong loading on one of the components, and no rating scale had a loading on both components. The second component saw strong loadings of extraversion and novelty-preference, which was in line with both the Overall PCA and the PCAs carried out on the ratings for the action clips. The first component saw strong loadings of approachability, neuroticism, trustworthiness and warmth. The scale neuroticism had a strong negative loading on the first but not on the second component. This was in line with the findings from the Overall PCA but not with the PCAs carried out on the ratings for the action clips (see Table 21).

Table 21

Factor Loadings of Six Rating Scales According to Display Format for Emotion Clips

	PC1			PC2		
Rating scale	FL	PL	*M*	FL	PL	*M*
Approachability	**.97**	**.98**	**.98**	.19	.15	.17
Extraversion	.11	.09	.10	**.99**	**.98**	**.99**
Neuroticism	**−.77**	**−.74**	**−.76**	−.35	−.30	−.33
NP	.12	.13	.12	**.98**	**.98**	**.98**
Trustworthiness	**.92**	**.92**	**.92**	−.24	−.10	−.17
Warmth	**.95**	**.95**	**.95**	.19	.14	.17

Note. PC = principal component number. NP = novelty-preference, FL = full-light displays, PL = patch-light displays, *M* = mean loading across display type. Boldface indicates loadings stronger than .40.

A reliability measure of the individual scales in each respective component was carried out; for the valence$_E$ construct we used emotional stability (neuroticism complemented). This revealed strong inter-item reliability for both the valence$_E$ construct (full-light α = .92, patch-light α = .92) and the dominance$_E$ construct (full-light α = .99, patch-light α = .92).

Linear combinations of the rating data were carried out to construct new variables. As before, dominance was defined as the mean of extraversion and novelty-preference. Valence$_E$ was defined in the same way as the overall PCA; that is, the mean of approachability, trustworthiness, warmth and emotional stability (neuroticism complemented). As before, the construct variables correlated strongly with the factor scores they were designed to approximate (see Table 22). The valence and dominance constructs were not correlated (ρ = .14, p = .17).

Table 22

Correlations (ρ) between Component Factor Scores and Linear Combinations of Rating Scales for Emotion Clips

	Valence$_E$ construct			Dominance construct		
PC	FL	PL	*M*	FL	PL	*M*
1	.98***	.96***	.97***	.06	−.02	.02
2	.04	-.01	.01	.98**	.98**	.98***

Note. PC = principal component number, FL = full-light displays, PL = patch-light displays, *M* = Mean coefficient across display format. *** $p < .001$.

Explanations of the Inconsistent Loading of Neuroticism

Recall that the same rating scales and the same paradigm were used for the action and emotion displays. The findings from the PCAs presented thus far showed that the main difference in loading patterns between the action and emotion clips was related to neuroticism ratings. For the valence construct, neuroticism ratings had a stronger (negative) loading for the emotion clips than for the action clips. For the dominance construct, neuroticism had a strong (positive) loading for action but for not emotion clips. Two possible interpretations for these findings are offered.

One possible interpretation of the different loading patterns on the valence construct could be that aspects of the movement associated with neuroticism differ depending on the behaviour. The anchors used to measure perceived neuroticism were *calm - anxious*. It could be that more exaggerated movement during a *star jump*, say, is not indicative of anxiousness, whereas for the emotion *sadness*, say, it is. This interpretation is supported by the finding that neuroticism had a strong loading on the dominance construct for action clips but not for emotion clips.

An alternative explanation is that calmness is perceived intrinsically as less attractive when an action is performed than when someone is displaying an emotion. That is, to be calm when portraying sadness, say, is a ‘good thing’ but to be calm when performing a star jump is not. This explanation fits with the findings from Chapters II and III, which showed that higher quantity of motion was associated with positive traits for actions but with negative traits for emotions; this was especially the case for the negative emotions *anger* and *sadness*.

Motion analyses. To test the first of the interpretations further we analysed the new construct and component variables in relation to the motion parameters Quantity of motion (QoM) and Contraction/Expansion index (CI; see Movement Analyses under Chapter II). If the aspects of movement that drive dominance and valence judgements on action clips are different from those on emotion clips, this should manifest in different relations between the motion parameters and the trait judgements depending on whether the display was an action or an emotion.

Using the valence and dominance constructs, calculated by combining the trait scales, the motion parameters were found to have a different relation with trait scales depending on the type of behaviour (action vs. emotion). Patterns in the trait ratings were similar across behaviour type (action vs. emotion clips) firstly in that there were significant positive correlations between CI and dominance whilst no correlations were found between CI and valence and, secondly, for both behaviours there were significant positive correlations between QoM and dominance (see Figure 9 and Figure 10). However, a different pattern between the behaviours was also found: QoM had a significant positive correlation with valence for action clips (ρ = .43) but not for emotion clips (ρ = .01). This confirms the earlier conclusion that exaggerated movement places someone in a favourable light when someone is portraying an action but not so when someone is portraying an emotion.

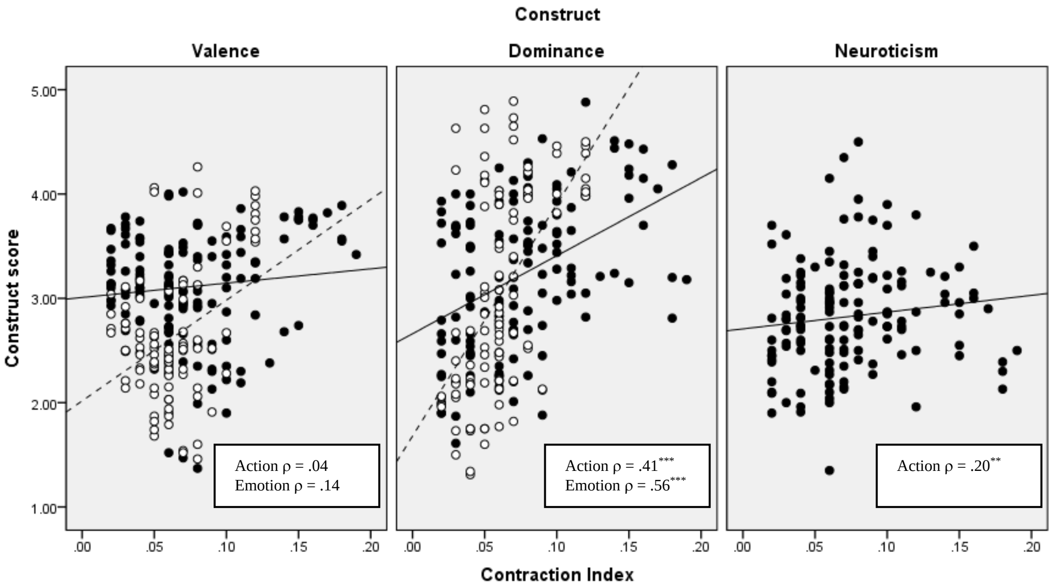


Figure 9. Construct scores against Contraction Index for action clips (solid markers/line) and emotion clips (nonsolid markers; stapled line). The valence construct for emotion clips includes neuroticism (complemented) whereas for action clips it does not. ** $p < .01$; *** $p < .001$.

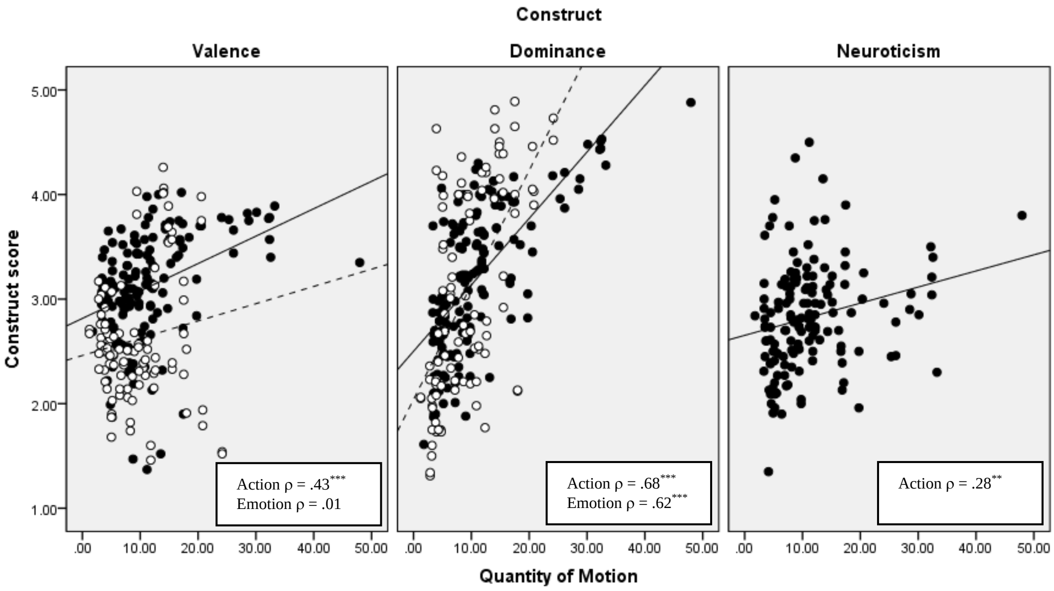


Figure 10. Construct scores against QoM for action clips (solid markers/line) and emotion clips (nonsolid markers; stapled line). The valence construct for emotion clips includes neuroticism (complemented) whereas for action clips it does not. ** $p < .01$; *** $p < .001$.

Experiment 3.2: Attractiveness judgements. To see whether the unequal pattern of the PCA was due to a different interpretation of the neuroticism scale depending on the displayed behaviour, another rating study was carried out. Here, we collected attractiveness judgements on a subset of the full-light stimuli. A differing correlation (direction or strength) depending on whether an action or emotion was displayed would suggest that the scale calm-anxious is interpreted differently by the observer depending on this display type.

Method. Twenty-two undergraduate psychology students (fives male and 17 females) took part in this rating experiment (mean age 21.05 years, *SD* = 7.63). All participants, who provided written, informed consent, took part in exchange for partial course credit.

A total of 150 full-light movie clips were used; half were action clips and half emotion clips. The clips included 57 of the 92 action clips used in the experiments reported in Chapter II and 38 of the 50 emotion clips used in the experiments reported in Chapter III. Participants were asked, for each clip, ‘how attractive do you find this person based on this movement?’, and were required to respond using a rating scale of 1 (very unattractive) to 5 (very attractive). Clips were presented in one block and were randomised for each participant. Further aspects of stimuli and procedure were identical to those described in Chapter II under Experiment 1.1.

Results and Discussion. High inter-rater reliability was found (α = .87) indicating that observers agreed with each other in which clips looked attractive and unattractive. Agreement was higher within emotion clips (α = .89) than it was within action clips (α = .67). To test whether these reliability coefficients were inflated by differences between the displayed types of behaviour (i.e. the different actions and the different emotional expressions), reliability analyses were carried out on all levels of the behaviour variables. In light of small sample sizes reliability coefficients were acceptable; apart from the actions touch toes (α = .02) and push (α = .50) and the emotion fear (α = .23), all behaviour types showed moderately strong reliability (all remaining αs > .60).

An independent samples *t*-test showed that clips were rated higher on attractiveness if they displayed an action (M = 2.64, SD = 0.35) than an emotion (M = 2.08, SD = 0.58), $t(148) = 7.22$, $p < .001$, $\eta^2 = .69$. One-way ANOVAs with type of display as fixed-factor were carried out for emotion clips and action clips separately, using a within-subjects design. These showed that the type of display had an impact on perceived attractiveness for both action clips, $F(7, 67) = 2.42$, $p = .03$, and emotion clips, $F(4, 70) = 81.19$, $p < .001$, although effect sizes were much higher for emotion clips ($\eta^2 = .82$) than for action clips ($\eta^2 = .25$). See Figure 11 and Figure 12.

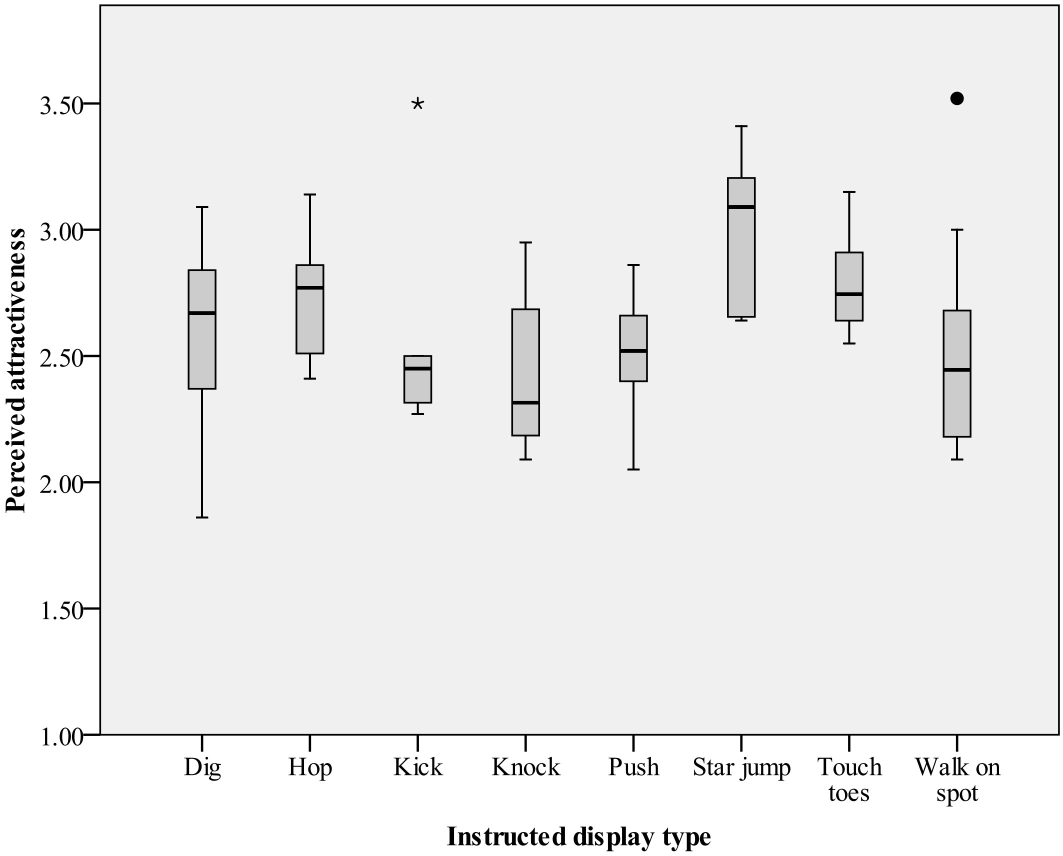


Figure 11. Perceived physical attractiveness of action clips.

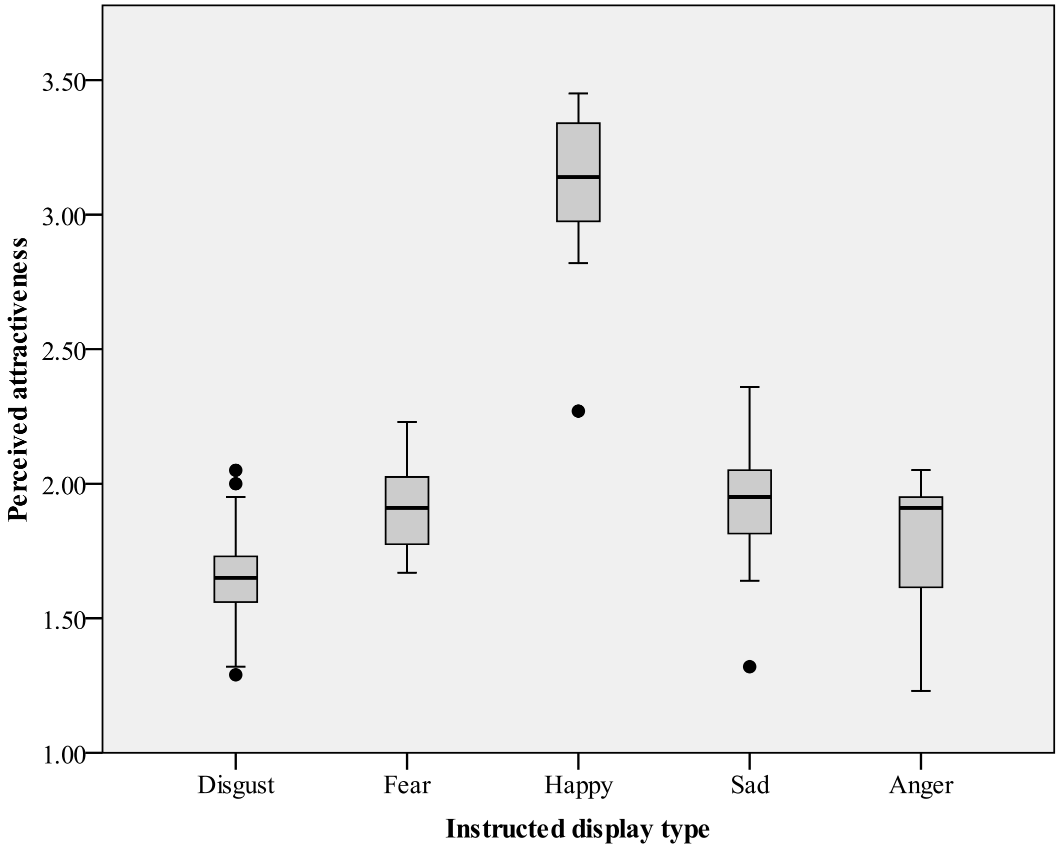


Figure 12. Perceived physical attractiveness of emotion clips.

The findings further showed that, as expected, dominance was strongly correlated with attractiveness for action clips (ρ = .68) but not for emotion clips (ρ = .22) (see Figure 13). This is in line with the unequal patterns found within the correlations between the motion parameters and the personality constructs: when a target displays more movement, this leads to positive judgements when the intended behaviour is an action but not when the intended behaviour is an emotion. However, there is also support for the second explanation: neuroticism was negatively correlated with perceived physical attractiveness for emotion clips (ρ = −.48) but also positively correlated with attractiveness for action clips (ρ = .29). This means that the unequal pattern from the PCAs may, in part, be due to a different interpretation of the scale calm-anxious.

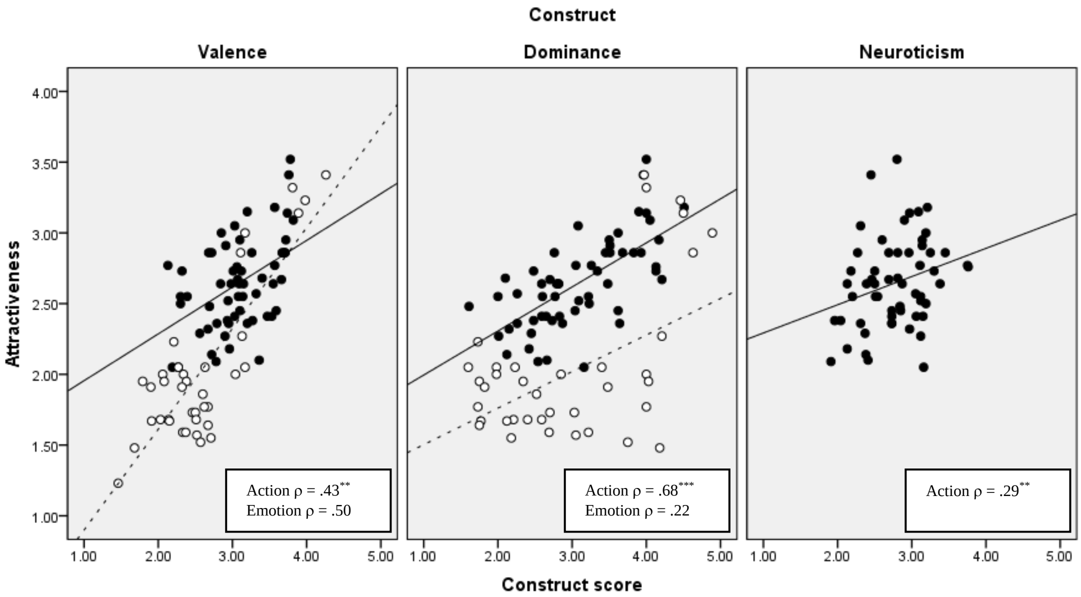


Figure 13. Attractiveness ratings against construct scores for action clips (solid markers/line) and emotion clips (non-solid markers; stapled line). The valence construct for emotion clips includes neuroticism (complemented) whereas for action clips it does not. Note: ** $p < .01$; *** p < .001.

The positive correlation between attractiveness and neuroticism found for action clips is not consistent with the loading pattern for this behaviour type (see Table 16). That is, because neuroticism had a negative loading on PC1 for action clips, it was expected to be associated with lower valence scores and this should manifest in a *negative* correlation with attractiveness. However, although the correlation was in the opposite direction to what was expected, it did not survive a Bonferroni correction for multiple comparisons, and this is further support for the finding that neuroticism does not correspond with the same trait judgements for emotion clips as for action clips.

Correlations between factor scores and attractiveness ratings were consistent with the pattern found for the constructs that approximated them; that is, for action clips attractiveness judgements were more strongly correlated with PC2 ($\rho = .52^{**}$)

than with PC1 ($\rho = .34^{**}$), whereas for emotion clips it was stronger for PC1 ($\rho = .52^{**}$) than for PC2 ($\rho = .22$, NS). This is further support for the adequacy of using the construct scores derived directly from the rating data.

The correlation between valence and attractiveness is not unexpected, but the coefficients were not perfect, which means that our observers' valence judgements were not solely driven by the 'what is beautiful is good' bias (Dion, et al., 1972) or, if they were, the scale *unattractive-attractive* is not enough on its own to tap into this bias. We therefore decided that subsequent rating experiments should continue collecting trait ratings on all scales underlying the construct.

General Discussion

The results indicate that the trait scales approachability, extraversion, neuroticism, novelty-preference, trustworthiness and warmth can be accounted for by two underlying components, which we labelled valence and dominance, respectively (after Oosterhof & Todorov, 2008). The components are by definition orthogonal, and constructs calculated from the underlying scales were only weakly correlated. This held independently of display format: PCAs on full-light stimuli gave the same patterns as PCAs on patch-light stimuli. The patterns revealed in our findings are consistent with research on trait impressions of faces (Oosterhof & Todorov, 2008; Todorov, et al., 2008). Thus, it appears that explicit trait impressions made on nonverbal stimuli by unacquainted observers might be reduced to judgements of two factors also for dynamic body stimuli.

However, the loading patterns varied between the behaviour types: when the displayed behaviour was an action, the trait scale neuroticism appeared to have loadings on both components, including a positive loading on the dominance component. We propose that this is partly due to different aspects of the movement driving neuroticism ratings based on the type of behaviour, as confirmed by unequal correlations between this trait scale and the motion parameters depending on the displayed behaviour. The unequal patterns emerging from PCAs on the action versus emotion clips might not be surprising, considering that behaviour in general influences perceived personality (e.g., Uleman, 2005) and furthermore we found that type of action (in Chapter II) and type of emotion (in Chapter III) had an impact on trait ratings using the stimuli.

It also appears that the differing loading patterns may be due to different interpretations of the scale *calm-anxious.* This was confirmed through collecting judgements of physical attractiveness, which were found to be positively associated with Neuroticism for action clips but not for emotion clips. We conclude that observers may indeed make only two judgements when asked to rate targets shown in full- or patch-light displays on our six rating scales: the anchors *calm* and *anxious* may simply be interpreted differently when an *action* is performed. The overall PCA, which combined both action and emotion clips, yielded a two-component solution similar to that of the emotion clips.

The findings were in line with research on person-perception (Dion, et al., 1972; Fiske, et al., 2007; Oosterhof & Todorov, 2008; Penton-Voak, 2009; Rosenberg, et al., 1968; Uleman, 2005). The idea of a small number of factors in decision making is, however, not novel, nor is it bound to person perception: using the principal of semantic differential (see Osgood, 1952), bipolar scales show that three factors apply to a multitude of decision making processes: these are typically labelled Evaluation, Potency and Activation (see Osgood, Suci, & Tannenbaum, 1957). The first two of these may be said to correspond to PC1 and PC2 identified in our data.

There is also evidence that the two-dimensionality of perceived personality has some basis in self-reports. Digman (1997), in an extensive meta-analysis of studies employing the five-factor personality model—including teacher-ratings, peer-ratings, self-ratings and inventories—suggested a two-factor model of personality. He labelled the first of his two factors α, which he proposed was associated with being well socialised, someone who has developed ‘impulse constraint and conscience, and the reduction of hostility, aggression and neurotic defense’ (pp. 1249-50). This may correspond to the first principal component identified in our PCAs, or its associated construct valence. Digman suggested a second factor (β) corresponding to ‘actualisation of self’ (p. 1250). Our second component, or the dominance construct, might somehow be linked to this factor: someone who strives for personal growth might show motion patterns that lead to high ratings of extraversion and novelty-preference.

As is the case for any principal component analysis, the labels for our two components are debatable. Nonetheless, as shown, there is evidence that the findings from the principal components analysis are ‘interpretable’, they are in line with most other research on person perception, and may have some foundation in personality traits as assessed through self-ratings and ratings by known peers.

All rating studies presented thus far showed high inter-rater reliability, in line with the much-documented consensus at zero acquaintance (Albright, et al., 1988; Albright, et al., 1997; Ambady & Rosenthal, 1993; Riggio, et al., 1990). The finding that our judgements could be reduced to two components and that these appear to be in accordance with evidence from self-ratings of personality leads to the question as to whether there is some degree of accuracy in people’s judgements, despite the apparent Halo effect, a bias that often begets erroneous personality judgements (Dion, et al., 1972; Feingold, 1992). This possibility is treated further in Chapter V.

PART 2:
POINT-LIGHT DISPLAYS OF TARGET WALKERS

Chapter V: Analysis of Individual Differences in Gait Based on Point-light Walkers

In the experiments presented thus far, we have shown that people make trait impressions based on exposure to full-light and patch-light displays of whole-body movements depicting target individuals engaging in a given action (Chapter II) or emotion (Chapter III). In four experiments we showed that inter-rater agreement was high amongst observers of these targets. Half the experiments employed full-light stimuli (whole silhouette of target visible) or patch-light stimuli (only patches were visible, but some static information was still available to observers). Patch-light stimuli were used because they contain less static information than the full-light stimuli. The finding that patterns in personality judgements were similar across the two display formats suggests that some dynamic aspects of the stimuli affect personality judgements even when static information is available to observers. However, some caution must be drawn because the patches still contain static information through the size and shape of the patches, potentially providing some static cues to the observer.

In the current and following chapters, we used point-light stimuli in which target individuals were instructed to walk naturally. Point-light displays constitute a popular form of impoverished visual stimuli to allow investigation of the contribution of motion (kinematic and form-from-motion) cues to observer ratings (Johansson, 1973; Troje, Westhoff, & Lavrov, 2005; Vanrie & Verfaillie, 2004). Information about the static form of the body is greatly reduced or even eliminated.

This chapter presents the stimuli used for Part 2 of the thesis. Ostensibly, the chapter looks at all aspects intrinsic to the walkers; that is, physical features of the stimuli as well as personality traits of the walkers themselves. Perceived personality is left for Chapter VI and VII.

First, the rationale for the choice of stimuli and the methods for selecting and capturing motion data are described. We then perform a motion analysis of the point-light walkers (Experiment 4.1): this relied on the use of Principal Component Analysis (PCA) to reduce the dimensionality of the motion data prior to extracting motion parameters describing a small number of components, following the procedure of Troje (2002). Motion parameters are further labelled by gathering subjective ratings of the motion characteristics from a new group of observers (Experiment 4.2). A description of the creation of the final point-light stimuli based on the motion data is also provided. Finally, in Experiment 4.3, the link between gait and self-reported personality is assessed, through comparing scores from personality questionnaires of target individuals with motion parameters extracted from their motion capture data.

Selection of New Stimuli

Display type. One methodological issue concerning the stimuli used thus far is due to the physical characteristics of the types of stimuli, i.e. full-light and patch-light displays. The video clips may have contained static information that in itself influenced trait ratings, and which were also related to motion parameters. For both display types, steps were taken to make sure the stimuli used contained visual information about the body only; for instance, many factors that bear an impact on trait ratings were absent, such as facial expressions (Montepare & Dobish, 2003), attractiveness (Dion, et al., 1972; Feingold, 1992; B. C. Jones, Little, Burt, & Perrett, 2004), and attire or well-groomedness (Albright, et al., 1988).

However, the full-light stimuli still contain a lot of static information. Full-light stimuli reveal body shape, and a bias against obesity is well documented. Overweight individuals are judged as lazy and incompetent (see Puhl & Brownell, 2001 for a review). This link could also be mediated by perceived attractiveness. Perceived attractiveness, in turn, is correlated with actual body mass index (BMI) of targets (Swami & Tovee, 2005), volume to height ratio (VTH) (Fan, et al., 2005; Fan, et al., 2004) and leg length relative to overall height (Swami, Einon, & Furnham, 2006). Furthermore, target gender was easily detectable from the stimuli and this may have driven the trait impressions directly. For example, female targets used more of both personal space and general space than did male targets (as measured with the EyesWeb CI and QoM indices, respectively), and females were also rated higher on warmth than were males.

Whilst the patch-light stimuli do remove much of the static information available to observers, some limited information about body shape is still available in these stimuli. The patches' size was directly linked to limb circumference (see Figure 1 in Chapter I). We therefore created new stimuli using point-light displays (Johansson, 1973). We used three-dimensional motion data recording for the creation of the point-light displays whereas stimuli used thus far contained two-dimensional data points. This therefore allowed for richer kinematic data for the purpose of studying motion parameters. The motion analyses are presented in Experiment 4.1.

Movement type. The instructions during the creation of the stimuli presented so far were found to influence perceived personality. It is therefore possible that observers use first-step decisions related to recognising the type of display to drive subsequent decisions of personality. We found, in Chapter II, that type of action influences perceived personality. For instance, star-jump stimuli were judged to be more extraverted than were other actions. We also showed that trait impressions are related to expressed emotion; e.g., displays of fear were considered neurotic. In the following experiments we wanted to know whether consensus at zero acquaintance could be found for neutral movements (i.e. in the absence of any instructions) using a singular, relatively uniform, everyday movement. We wanted to use a movement high in ecological validity in terms of occurrence in the real world, and chose to create stimuli of target individuals walking naturally. Walking is also a movement that reflects individual differences (Loula, et al., 2005), as further evidenced by the observation that one easily can recognise individuals from their gait (Cutting & Kozlowski, 1977). Finally, walking occurs in situations where several observers may be scrutinising the movements of a person (e.g., in a job interview or an audition).

Creation of New Stimuli

Target individuals. Twenty-six targets (14 females and 12 males; mean age = 19.7 years, SD = 2.6) were recruited from Durham University. Apart from three males, all participants were undergraduate psychology students who took part in the experiment in exchange for partial course credit. A selection criterion ensured that targets had little, or no, acting experience. A range in the target individuals' personality traits was ensured through a selection procedure whereby prospective targets filled in the NEO FFI SF personality questionnaire (Costa & McCrae, 1992). Ranges were approximately 29^{th} to 72^{nd} percentiles in all five constructs of the Big Five. Due to the specificity of the population as well as the recruitment procedures, the targets' traits were not in perfect accordance to the general population. Notably, one-sampled t-tests comparing our targets' scores to the population mean, showed that they scored low on conscientiousness (t_{25} = 3.03, p < .01), and high on extraversion (t_{25} = 2.98, p < .01) and openness to experience (t_{25} = 4.74, p < .001).

Male and female targets did not differ significantly in traits as measured through the standardised scores, apart from on agreeableness, with females scoring higher (M = 55.6, SD = 5.61) than males (M = 42.4, SD = 13.46; t_{24} = 3.34, p < .001). See *Figure 14*.

Because of the potential confound of perceived Body Mass Index (BMI) and trait ratings we collected BMI data on 21 of the target individuals. BMIs were within the normal range (M = 21.3, SD = 2.88) and no significant difference between males and females was found (t_{21} = 1.71, p = .10).

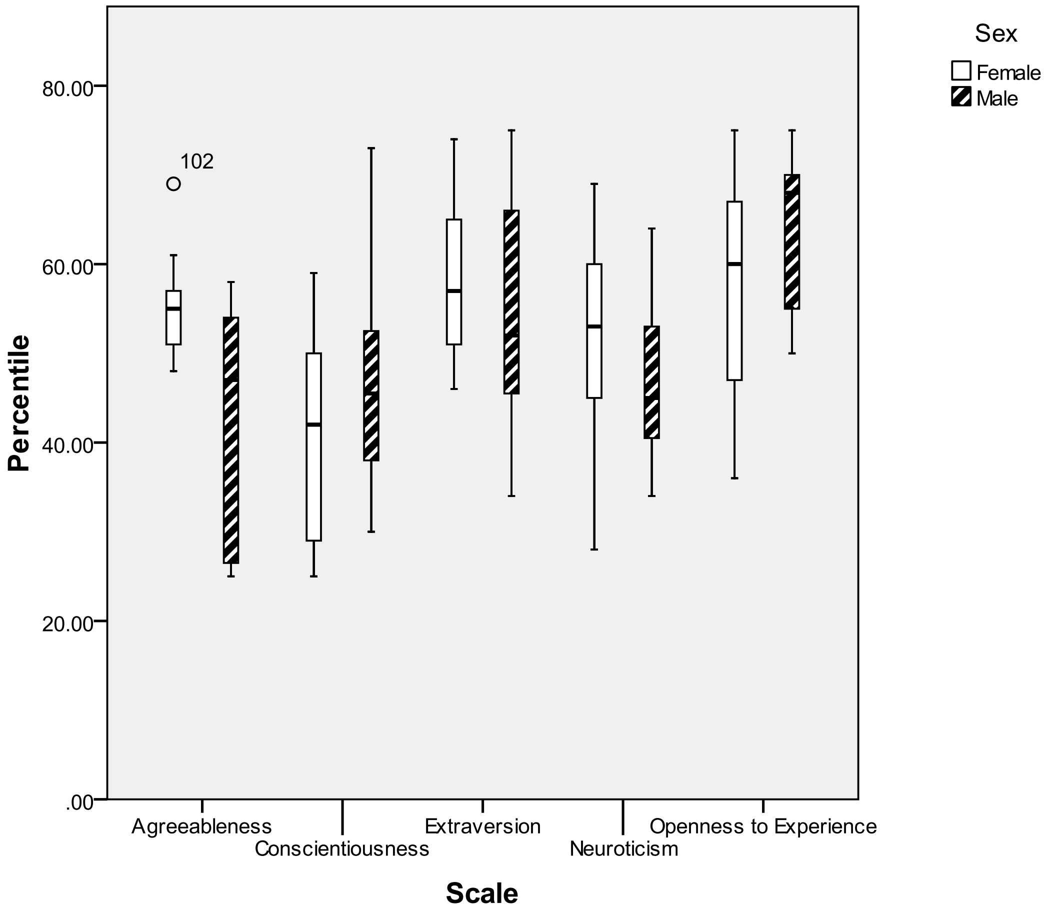


Figure 14. Self-reported personality scores from the NEO FFI SF. Scores represent a respondent's percentile in relation to the general, target-specific population.

The motion capture procedure. The target individuals' whole-body walking movements were captured using a VICON system (Oxford Metrics, UK). Eighteen retro-reflective markers were attached to targets' joints at the targets' feet, knees, hips (four markers), torso, shoulders, head, elbows, wrists and hands. Three-dimensional positions of these markers were recorded at a frequency of 100 fps.

Targets were instructed to walk at their own desired pace between two spots approximately eight metres apart. We selected one whole walk cycle from the middle of this sequence. The starting point was defined by one foot touching the floor; cycles were selected such that they could be looped continuously without looking 'jerky'. However, to avoid the possibility that some stimuli looked smoother than others, a 150 ms black frame was added to the end of each clip. Since targets' speed differed, the number of frames was not equal (mean number of frames = 114.4, *SD* = 11.2).

Video conversion. The final step of creating point-light stimuli consists of converting capture data to two-dimensional video format. This was done using an in-house script in Matlab (The MathWorks Inc., Natick, MA, USA). To reduce the amount of static information available to observers, the number of markers in the final 2D stimuli was reduced to 13, a number commonly used in whole-body point-light displays (Dekeyser, Verfaillie, & Vanrie, 2002; Loula, et al., 2005; Prasad & Shiffrar, 2009). This involved averaging the two left and two right hip markers to create a single virtual left and a single virtual right hip marker. This meant that the placement of the hip markers was to a degree standardised: the markers were 'inside' the hips of the targets, thus reducing variability of perceivable waist circumference. For the other 11 markers selected, one was from the target's head, and one from each shoulder, elbow, wrist, knee and foot. In the final point-light video clips, translation was removed and the targets appeared as if walking on a treadmill, facing diagonally towards the right in three-quarter profile view. Three-quarter profile was chosen over frontal, half-profile or profile view because the patch-light and full-light displays used in Part 1 were frontal with minor deviation from this frontal view and we wanted the final stimuli to mimic this. Because size was kept constant and because limbs never obstruct the point-lights, an observer could not tell whether the walker was walking towards or away from the observer. However, it is likely that observers will report walkers as approaching them, due to a facing bias (Vanrie, Dekeyser, & Verfaillie, 2004). The points were white on a black background (see Figure 15). To avoid the possibility that some stimuli looked smoother than others when the walk cycle was repeated, a 150 ms black frame was added to the end of each clip.

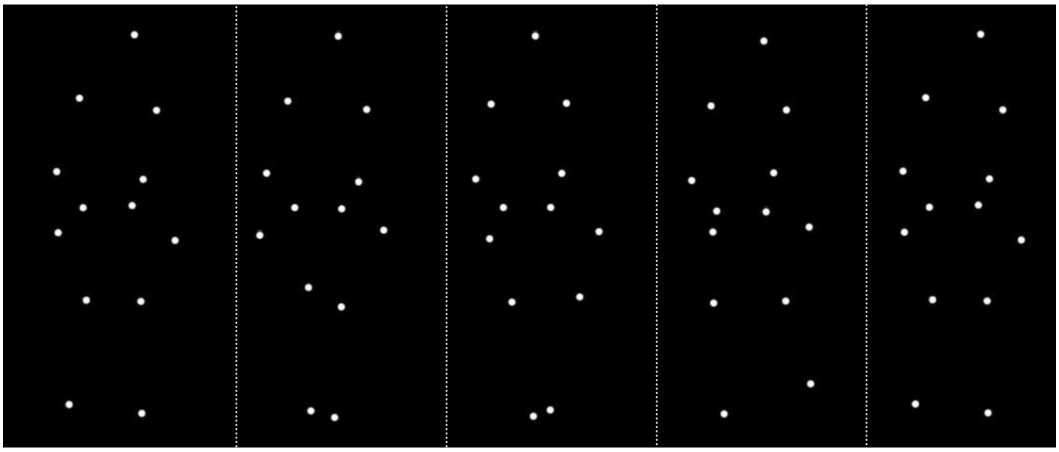


Figure 15. Snapshots from an example point-light walker sequence.

Experiment 4.1: Motion Parameters of Point-light Walkers

Method and Results. To extract motion parameters from the point-light walkers, we chose the methods outlined by Troje (2002). A range of motion parameters was extracted from the point-light walker stimuli. The walker stimuli were analysed individually: each of the 26 motion data files used to generate the point-light stimuli was run through a separate Principal Component Analysis (PCA), using Matlab. The data was standardized by subtracting the average posture before being analysed: the average Cartesian coordinates of the markers over the walk cycles, effectively placing each walker on a 'virtual treadmill'. Only the 13 markers that were used to create the point-light stimuli were retained, to limit the possibility of extracting parameters that were unavailable to observers.

The 3D coordinates of each marker over time were used in the PCA, thus resulting in 39 variables, each line corresponding to one frame, following the procedure of Troje (2002). The number of data points thus equalled number of frames required by the walker to complete one cycle. The PCAs yielded an eigenposture, the weights, and a mean posture, for each of the 52 motion-capture files. Using a selection criteria of eigenvalue > 1, all the PCAs yielded a three-factor solution (average eigenvalues: $eig_{PC1} = 34.2$; $eig_{PC2} = 2.6$, $eig_{PC3} = 1.7$). The analyses further showed that, averaged across all walker files, two components explained 94.3% of the data (of which the first accounted for 87.7%), and including a third component explained 98.7%.

As was the case in Troje's (2002) study, the scores of the principal components followed a sinusoidal curve. Using the Curve Fitting Toolbox in Matlab we identified the amplitude, frequency and phase of three components for each walker file. The goodness of fit for all three components was strong (across motion-file average $R^2_{PC1} = .99$; $R^2_{PC2} = .92$; $R^2_{PC3} = .91$). The third component accounts for only 4.4% of the variance and its discernible impact on the movement was minimal[3], this component was therefore not included in further analyses. Figure 16 shows the scores and sinusoidal fit for the first two PCs of an example walker.

[3] This discernible impact was based on visual inspection of motion data files created by exaggerating the components (see Experiments 5.3.1 and 5.3.2 in Chapter VI).

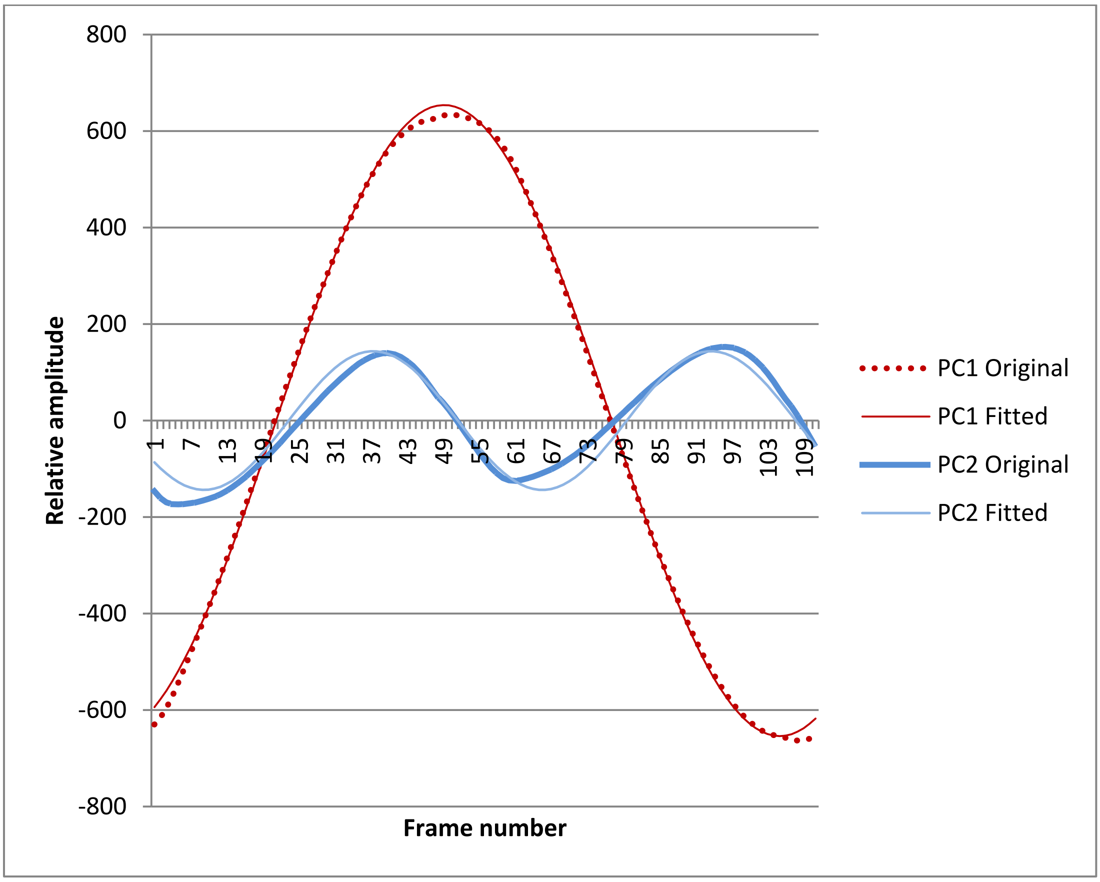


Figure 16. Scores from two components (stapled) as well as sinusoidal fits (solid) of an example walker. For this walker R^2 for PC1 >.99, R^2 for PC2 = .94.

It was expected that similar patterns would emerge for all the 52 PCA analyses; i.e. that the variables have a consistent loading on the respective components across all walker files. However, since the PCAs were run separately, it is likely that some components would be negated due to the arbitrariness intrinsic to principal component analyses; as an analogy, someone's *height in cm* may have a positive or negative loading on the one and same construct but depending on the sign this would be labelled 'shortness' or 'tallness'. Therefore, a procedure of negating those loadings that appeared to be inverse of the 'typical' pattern was performed. To do this, we calculated the correlation between one column of loadings for a given component *i* for a given motion capture file *j* (part of set *S*) with the mean loadings of the same component *i* for all other motion capture files (*S* excluding *i*) and negated all those with a correlation below -.50. A correlation strength threshold of .50 was chosen as to minimise the possibility of artificially inflating the model's goodness of fit. The resulting matrices were now equalised so that all the components could be interpreted in the same direction.

Once this was done for the first four components of all motion capture files, we checked for consistency of loadings using the Spearman-Brown formula to calculate Rosenthal's reliability coefficient, *R* (Rosenthal, 1973). This showed excellent reliability for PC1 (R = .99), PC2 (R = .87) and PC3 (R = .96).

No link was found between the BMI of the targets and the motion parameters (all absolute *R*s < .35, all *p*s > .11). This is suggestive that the motion parameters are indeed indicative of dynamic, and not static, information.

Inspection of the weightings of the motion data variables on these components yielded no intuitive description of the components (e.g., markers close in location, such as wrists and hands, were not grouped) and giving a name to the components was therefore difficult using loading patterns. Component 1 consistently repeated itself after one walk cycle and thus represents the fundamental frequency of the walker; correspondingly, this parameter was strongly negatively correlated with the number of frames required to complete a cycle (r = -.88). The frequencies of PC2 and of PC3 were typically twice that of PC1.

Experiment 4.2: Effort-shape analysis. An alternative method to describe the motion parameters was performed by collecting subjective motion descriptions from a new group of observers (Experiment 4.2). To do so, we used an Effort-Shape analysis, which is an approach for describing movement, originating in works by Laban (Gross, Crane, & Fredrickson, 2010). Shape judgements included information on contraction or expansion of the body (do limbs move towards or away from the body centre; is the body gathering or scattering), which was collected through judgements of torso and limb, and Effort parameters, which described forcefulness during movement through judgements of Energy, Space, Time and Flow (Gross, et al., 2010).

Method and Results. A new group of observers (N = 26, 12 males and 14 females, mean age = 19.5 years, SD = 1.0) were used. Observers were shown the point-light walkers on a 19" LCD screen with a refresh rate of 60 Hz. Distance between participant and screen was not controlled but at an estimated average of 40 cm the stimuli a visual angle of 22º (vertical). PsyScope X (http://psy.ck.sissa.it) was used to present the stimuli and collect participant responses, using the top row number keys of a standard QWERTY keyboard. During a trial, a five-point rating scale was presented to the observers, together with anchors (see Table 23 for descriptions of the anchors used). Effort-shape components were blocked and block order was randomised, as was order of stimulus-presentation within the block.

Table 23 shows the correlations between the effort-shape descriptors and the sinusoidal parameters for the first two components. The Effort-Shape analysis differentiated PC1 from PC2 with regards to the amplitude only. PC1 Amplitude was associated with use of personal space, with expanding torso and limbs moving away from the body. PC2 Amplitude was negatively correlated with Time, Space and Flow; thus associated with a *leisurely*, *relaxed* walk with more *diffuse* use of space.

Table 23

Effort-Shape Descriptors and Correlations (ρ) with Component Amplitudes

Effort-Shape	Descriptors presented to observers		PC	
Factor	Left-anchor	Right-anchor	1	2
Torso	Contracted, bowed, shrinking	Expanded, stretched, growing	.52**	−.11
Limb	Moves close to body, contracted	Moves away from body, expanded	.40*	.05
Energy	Light, delicate, buoyant	Strong, forceful, powerful	.03	−.34
Space	Indirect, wandering, diffuse	Direct, focused, channelled	.20	−.54**
Time	Sustained, leisurely, slow	Sudden, hurried, fast	.30	−.52**
Flow	Free, relaxed, uncontrolled	Bound, tense, controlled	.18	−.67**

Note. PC = principal component number. * $p < .05$; ** $p < .01$; *** $p < .001$

General Discussion: Motion Parameters

In general it appears that PCA is appropriate for modelling the whole-body, single gait-cycle walking movements captured for the purposes of this study (see also Troje, 2002). The three-dimensional positions of 13 markers were reduced to two PCs, retaining 94% of the variance of the data, and these two components followed a sinusoidal model with a solid goodness of fit (R^2_{PC1} = .99 and R^2_{PC2} = .92). Since the third component accounted for only 4.4% of the variance of the motion data this component was not included in further analyses because its impact (discernible or otherwise) on the movement was minimal. The phase of a sinusoidal function is simply related to the arbitrary decision of where that walk cycle is started and which direction the walker was headed, and is therefore excluded from all further analyses. The final four parameters (amplitudes and frequencies of PC1 and PC2) were thus employed for analyses on personality, both self-reported (in this chapter) and perceived (in Chapters VI and VII).

Subjective descriptions of the movement were collected by presenting the stimuli to a separate group of observers. The findings indicate that the first principal component is associated with high use of personal space, corresponding to the motion parameter CI from Part 1 of the thesis. The second principal component seems to be associated with a leisurely, uncontrolled walk, with more diffuse use of space.

Experiment 4.3: Self-Reported Personality and Motion Parameters

The first question that may arise when considering the link between personality and whole-body motion is whether people with certain personality characteristics display different types of movement. As discussed in Chapters II and III, the belief that styles of expressive behaviour are linked to personality has been around for some time (Friedman, et al., 1980; Riggio & Friedman, 1986; Riggio, et al., 1990). Friedman, et al. (1980) found that expressive physicians are dominant and adventurous: this was assessed through audio and audio-video stimuli of neutral sentences by naive listeners. Physiological reactions to instructions have also been found to differ depending on personality traits (Bono & Vey, 2007).

Method and Results. Male and female target walkers were treated separately for the analyses. For the female walkers, there were no significant correlations between any of the self-reported Big Five personality traits and the sinusoidal parameters extracted from the motion-capture files (all *p*s > .05). Extraverted males appeared to have a reduced PC1 amplitude, $\rho = -.72$, $p = .009$. However, this did not survive a Bonferroni correction for multiple comparisons. Males who rated themselves open to experience had a markedly lower PC2 amplitude, $\rho = -.90$, $p < .001$. Although this correlation survived Bonferroni correction, the corresponding correlation for female target walkers was nil ($\rho = -.02$, $p = .93$) so it is unclear whether this is simply an artefact in the data. We note that the low number of walkers has resulted in limited power for this analysis and conclude thus that there may be a trend for male walkers; a follow-up study with a larger sample is needed to lower the probability of committing a Type II error.

Discussion. In summary, there appeared to be no link between motion parameters and self-reported personality. This was based on target individuals' scores on the NEO FFI SF personality questionnaire (Costa & McCrae, 1992) and motion analysis following the methods of Troje (2002).

General Discussion

In this chapter we have presented the point-light walkers: the stimuli employed for the second part of this thesis. The selection and capture procedures were justified and presented; we also showed how point-light walkers can be described by two sinusoidal functions. An assessment of the link between these extracted motion parameters and the self-reported personality traits of the target walkers was also performed.

No conclusive link was found between a target's self-reported personality and his or her gait. This may, at a first glance, seem unexpected, considering the research that suggests a link between expressiveness and body motion (Friedman, et al., 1980; Riggio & Friedman, 1986; Riggio, et al., 1990). However, it is possible that the physical settings as well as the instructions given to the walkers during the motion capturing created restrictions of personal expressions of emotions. It is possible that the situation in which the motion capture took place may have caused a stale walk, not representative of motion that may occur during, say, a job interview. One must therefore be careful to conclude that personality is not reflected in gait.

In the subsequent chapters, we focus our attention back to perceived personality, based on the stimuli and motion parameters presented in this chapter. In Chapter VI, we establish consensus at zero acquaintance based on the point-light walkers. The kernel of truth hypothesis is investigated through continued use of the self-reported personality scores reported in this chapter. This assumes, of course, that there *is* indeed a link between gait and self-reported personality, as assessed through the NEO FFI, but that our motion analyses simply failed to highlight it. The personality ratings are further compared to the motion parameters extracted in this chapter, to find whether physical, dynamic aspects of the stimuli can explain personality trait ratings. Finally, the amplitudes of the first two components are scaled to see whether personality impression can be affected.

Chapter VI: Perceived Personality of Point-Light Walkers

In Chapter V, we presented the stimuli used for the second part of the thesis. We showed that motion data of target individuals walking naturally could be modelled by sinusoidal functions. The original 39 variables were reduced to two components using principal component analysis (PCA), explaining 94% of the variance of the data, and these two components were each summarised by two motion parameters: amplitude and frequency of the sinusoidal fit.

Using these four motion parameters we found no link between a person's gait and their self-reported personality, although some shortcomings of the setup were offered, notably related to methodological issues surrounding recording and analysis of motion data. These shortcoming, as well as others, are treated in more detail in Chapter VII. However, even if we were to conclude that there is no link between gait and a person's self-reported personality, this does not by any means indicate that gait is unrelated to how people are perceived. In the current chapter, we investigate the perception of personality of the stimuli presented in Chapter V.

First, in Experiment 5.1, we establish consensus at zero acquaintance based on the point-light walkers. The procedures of this experiment follow those used in experiments presented in Chapters II and III. Having established consensus at zero acquaintance in Experiment 5.1, we set out to find motion parameters that drive the trait ratings. To do this, we first, in Experiment 5.2, compare the personality ratings to the motion parameters identified in Chapter V. Two motion parameters, related to the first two principal components, are found to be related to perceived personality. To validate these findings, two further experiments are carried out, whereby the first (Experiment 5.3.1) and second (Experiment 5.3.2) principal components are manipulated to see whether this affects personality impressions.

Experiment 5.1: Personality Ratings on Unmodified Point-Light Walkers

Consensus at zero acquaintance is almost consistently reported in the literature for facial stimuli when people are asked to make personality judgements of strangers (Albright, et al., 1997; Berry, 1990; Oosterhof & Todorov, 2009; Penton-Voak, et al., 2006; Willis & Todorov, 2006). However, movement too has an impact on personality ratings, as shown by consensus at zero acquaintance for personality judgements made on dynamic body stimuli (e.g., Ambady & Rosenthal, 1993; Brownlow, et al., 1997; Kenny, et al., 1992). Agreement in personality trait ratings has also been found using point-light walker stimuli (Heberlein, et al., 2004; Montepare & Zebrowitz-McArthur, 1988). The present experiment was carried out in order to, firstly, validate these findings by establishing consensus at zero acquaintance based on our point-light walker stimuli and, secondly, to find underlying motion parameters that can explain personality trait judgements.

Method. Twenty-four new observers were used (11 males and 13 females; mean age = 19.3 years, *SD* = 0.9). All participants were undergraduate psychology students who took part in the experiment in exchange for partial course credit. The participants were from different year groups than that of the target walkers to avoid familiarity.

As in the previous experiments (e.g., Experiment 1.1, Chapter II), observers rated the stimuli on six rating scales: approachability, extraversion, neuroticism, novelty-preference, trustworthiness and warmth/friendliness. PsyScope X (see http://psy.ck.sissa.it) was used to present stimulus movie clips on a 19" LCD screen with a refresh rate of 60 Hz. Participant responses were collected using the top row number keys of a standard QWERTY keyboard. The personality scales were blocked by trait scale. Block order was randomised, as was order of stimuli within the block. All further aspects of the Method were identical to those used in previous chapters.

Results and Discussion. Good overall reliability was found, with Cronbach's alpha for all stimuli combined (across walk-type and sex) on all scales above .83 (see

Table **24**). Overall agreement was not inflated by sex of target, as within-sex agreement too was high.

Table 24

Inter-Rater Reliability (α) of Personality Ratings Based on Unmodified Point-Light Walkers

	Target sex		
Rating Scale	F	M	All
Approachability	.74	.85	.81
Extraversion	.92	.90	.91
Neuroticism	.67	.78	.87
Novelty-pref.	.94	.93	.93
Trustworthiness	.79	.79	.78
Warmth	.74	.82	.78

Note. Novelty-pref. = novelty-preference, F = female, M = male. Column 'All' indicates reliability coefficients obtained when all targets are pooled.

The high inter-rater reliability fits with previous studies showing consensus at zero acquaintance (Albright, et al., 1988; Engell, Haxby, & Todorov, 2007; Kenny, Albright, Malloy, & Kashy, 1994) and shows that some physical features of the stimuli, visible to observers, drive trait impressions.

Experiment 5.2: Motion Parameters associated with Perceived Personality

Consensus at zero acquaintance based on visual stimuli alone means that discernible visual characteristics communicate trait impressions and it should thus be possible to extract these. Some studies have indeed succeeded in identifying static visual cues for trait judgements, such as facial symmetry in photographs (Noor &

Evans, 2003) or low spatial frequencies (Bar, et al., 2006). However, describing purely *dynamic* cues for trait ratings has proven more difficult. Ambady and Rosenthal (1993) found agreement amongst judges of personality traits based on 30 s silent video clips. Interactions with the environment (‘fidgeting’ with hands or objects) as well as facial expressions (smiling and frowning) were the only parameters that predicted personality or teacher effectiveness ratings. Montepare and Zebrowitz-McArthur (1988) found high reliability for personality trait ratings of point-light walkers (e.g., dominance, boldness and approachability), despite any evidence of accuracy in the trait ratings. They identified no motion parameters predicting these personality ratings. Rather, they used subjective gait ratings, which may be confounded with trait ratings. High inter-trait scale correlations are common in rating studies (e.g., Oosterhof & Todorov, 2008), and a ‘what is beautiful is good’ bias, the so-called Halo Effect, is common when making judgements of strangers in general (Dion, et al., 1972).

In contrast to studies examining relations between trait ratings alone, we aimed to discover which objective aspects of the physical stimulus – specifically, which visual cues specified in the kinematics of people’s gait – drive personality trait judgements. Although kinematic analyses of gait and other whole-body movement have been used to discover which visual cues drive perception of sex (Kozlowski & Cutting, 1977; Mather & Murdoch, 1994; Troje, 2002), vulnerability (Gunns, et al., 2002; Johnston, et al., 2004) or emotion (Pollick, Paterson, et al., 2001; Roether, et al., 2009), we here present a kinematic analysis of personality trait judgements.

Method and Results. In order to identify motion parameters associated with personality ratings we correlated trait ratings obtained in Experiment 5.1 with motion parameters previously extracted from the motion data (see Chapter V).

The amplitudes of PC1 and PC2 showed potential in predicting trait impressions (see Table 25); correlations between trait impressions and other components were nil (data not reported). The amplitude of PC1 was correlated with extraversion, novelty-preference, trustworthiness and warmth. The amplitude of PC2 correlated negatively with neuroticism ratings (p = .018) although this did not survive a Bonferroni correction for multiple comparisons.

Table 25

Correlations (ρ) between Component Amplitudes and Personality Ratings

	PC	
Rating scale	1	2
Approachability	.22	.29
Extraversion	$.50^{**}$	−.18
Neuroticism	.07	$-.56^{*}$
Novelty-preference	$.59^{**}$	−.27
Trustworthiness	$.63^{***}$	.01
Warmth	$.62^{***}$	.12

Note. PC = principal component number. * p < .05; ** p < .01; *** p < .001.

Discussion. In Chapter V, where walk cycles were modelled sinusoidally, we showed how three-dimensional coordinates of 13 markers could be reduced to two components explaining 94% of the variance. These were further reduced to just four sinusoidal parameters. Out of these, two (the amplitudes of the first two principal component scores) were linked to how observers judge the personality traits of the walkers. PC1 amplitude was associated with perceived extraversion, novelty-preference, trustworthiness and warmth. PC2 amplitude was negatively associated with neuroticism. The findings add to the literature suggesting that there is a link between a person's gait and their perceived personality (e.g., Montepare & Zebrowitz-McArthur, 1988).

Since point-light stimuli are considered to reduce static information from the moving object, this finding is indicative of there being something about a walker's *movement* that drives trait impressions. That said, correlational studies are always open to criticism of confounding variables, and ours is no exception. Point-light displays could still contain some static information. Earlier methods for creating point-light walkers were more vulnerable to this criticism: the walker stimuli used by Kozlowski and Cutting (1977), for instance, allowed markers to be occluded by body parts of the target. The markers in our stimuli, however, were never occluded. Further to this, by using virtual hip markers that appeared to be within the core of the body, body weight status was harder to discern. Indeed, a separate rating study of 22 new observers (10 females, 12 males; mean age = 19.4 years, *SD* = 0.96) showed that there was no correlation between the BMI of the targets and their body weight status. Even though inter-rater agreement was high (α = .94), perceived body weight status was not associated with any personality trait ratings (data not reported).

However, other static variables may still be confounding the results. It is possible that ratios such as leg length to upper body height were visible and also affected the motion parameters. Instead of individually investigating all these potential confounds, we addressed these issues by carrying out two experiments where we manipulated the amplitude of the first (Experiment 5.3.1) or second (Experiment 5.3.2) component. These experiments also allowed us to explore whether existing motion capture data can be modified in a meaningful way to influence observers' perception. This is potentially of interest to those within the field of computer modelling of human avatars where the producers want to influence the audience's perception; e.g., computer games, commercials, cartoons and animated movies. We therefore wanted to ascertain whether it is possible to influence observers' trait impressions through manipulations of the PCs that are correlated with variation in trait ratings.

Experiments 5.3.1 and 5.3.2: Influencing Perceived Personality by Manipulating Point-Light Walkers

Motion data that were the basis for creating the point-light walker stimuli presented to observers in Experiment 5.2 were used for the manipulations. This was achieved by multiplying the amplitude of the scores of the first (Experiment 5.3.1) or the second (Experiment 5.3.2) principal components by given constants. Point-light movie clips were then recreated based on the first four components only, which were mapped onto the original three-dimensional positions of the markers, thus ensuring that differences between the versions of a given target walker's motion were due to motion alone. Although two components sufficed in explaining 94% of the variance in the data, we included the third and fourth component in order to make the walkers more 'natural-looking'.

Experiment 5.3.1: manipulation of PC1. In Experiment 5.3.1, we scaled the amplitude of the coefficients, and thus of the scores, of the first principal component (PC1) by –20% to +20% in 10% increments, thus creating 5 new versions of each point-light walker. A 0% scaling represents the original score for PC1.

Method. Twenty-six new observers (20 female and six male; mean age = 19.4 years, *SD* = 0.8) took part in this experiment. All participants were undergraduate psychology students taking part in exchange for partial course credit. Each block contained 130 stimuli, consisting of the five versions of each walker presented in a random order. Further details of the methods were identical to those in Experiment 1.

Results and Discussion. As can be seen in see Table 26, inter-rater reliability was high, with most coefficients suggesting strong agreement, with the exception of approachability (α = .66).

Table 26

Inter-Rater Reliability (α) of Personality Ratings across Five Levels of Manipulation of PC1

	Level of manipulation					
Rating Scale	−20%	−10%	0	+10%	+20%	Mean
Approachability	.74	.70	.62	.61	.64	.66
Extraversion	.89	.90	.89	.92	.94	.91
Neuroticism	.73	.79	.57	.77	.74	.72
Novelty-preference	.92	.90	.91	.92	.91	.91
Trustworthiness	.77	.79	.80	.74	.76	.77
Warmth	.78	.82	.74	.83	.80	.79

As can be seen in Figure 17, manipulation of PC1 had an impact on trait ratings. A multivariate analysis of variance (MANOVA) treating the 26 walkers as random effects confirmed this: There was a main effect of manipulation, $F(24, 332) = 14.58$, $p < .001$, $\eta_p^2 = .47$. This was significant for extraversion, $F(4, 100) = 127.61$, $p < .001$, $\eta_p^2 = .84$, novelty-preference, $F(4, 100) = 121.88$, $p < .001$, $\eta_p^2 = .83$, trustworthiness, $F(3.16, 79.04) = 4.80$, $p < .01$, $\eta_p^2 = .16$, and warmth, $F(3.35, 83.64) = 11.92$, $p < .001$, $\eta_p^2 = .32$. For all these scales, larger amplitudes resulted in higher trait ratings. As expected, there was no effect of PC1 on approachability (p = .60, η_p^2 = .03) or neuroticism (p = .56, η_p^2 = .03). This therefore coincides with the findings from Experiment 5.2. Trend analyses showed that there was a linear effect of manipulation on the ratings for all four trait scales: extraversion, $F(1, 25)$ = 300.79, $p < .001$, η_p^2 = .92; novelty-preference $F(1, 25)$ = 412.10, $p < .001$, η_p^2 = .94; trustworthiness, $F(1, 25)$ = 10.12, $p < .01$, η_p^2 = .29; and warmth, $F(1, 25)$ = 34.20, $p < .001$, η_p^2 = .58. There were no significant higher-order trends.

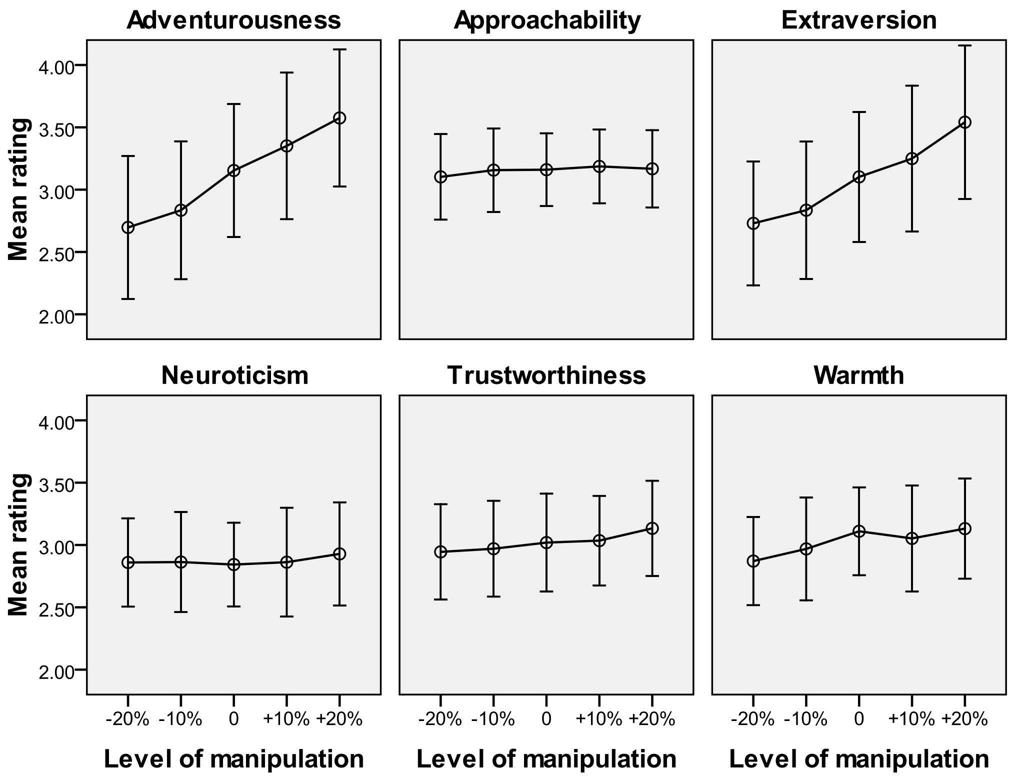


Figure 17. Mean trait ratings (± 1 *SD*) on six scales for five levels of manipulation of amplitude of PC1 scores.

Experiment 5.3.2: Manipulation of PC2. A similar procedure as in Experiment 5.3.2 was carried out for PC2; that is, the amplitude of the score was manipulated in order to see what effect this had on perceived personality traits. PC2 accounted for far less variance in the data than did PC1 (6.6% versus 87.7%) and, in line with this, manipulations had to be more pronounced in order to make a discernible difference to observers of the point-light movie clips. Initial visual inspection of the movie clips showed that exaggerating the amplitude of the scores of PC2 (i.e. manipulation > 0%) quickly resulted in non-natural-looking walkers compared to diminishing the amplitude by the same percentage. Thus, we scaled the amplitude of the scores of PC2 by −60%, −40%, −20%, 0%, and +20%, thus creating another set of 5 modified versions of each point-light walker.

Method. Twenty-one new observers were used (five male, 16 female), drawn from the same population as Experiment 5.3.1 and with similar age demographics (M = 21.1 years, SD = 2.9). All further procedures were identical to that of Experiment 5.3.1.

Results and Discussion. There was good inter-rater reliability amongst observers on most scales (neuroticism $\alpha = .66$; all other αs > .71), as shown in Table 27.

Table 27

Inter-Rater Reliability (α) of Personality Ratings across Five Levels of Manipulation of PC2

	Level of manipulation					
Rating Scale	−60%	−40%	−20%	0	+20%	Mean
Approachability	.74	.78	.78	.88	.70	.78
Extraversion	.93	.92	.93	.89	.92	.92
Neuroticism	.70	.75	.59	.51	.74	.66
Novelty-preference	.88	.89	.90	.86	.87	.88
Trustworthiness	.68	.66	.72	.87	.77	.74
Warmth	.62	.73	.71	.85	.66	.71

The findings show that manipulation of PC2 had an effect on trait ratings, $F(24, 332.6) = 2.09$, $p = .002$, although this effect was small ($\eta_p^2 = .12$). This held for novelty-preference alone, $F(4,100) = 6.45$, $p < .001$, $\eta_p^2 = .21$. As illustrated in Figure 18, trend analyses show that the only significant linear effect of amplitude manipulation was on novelty-preference, $F(1,25) = 10.49$, $p = .003$, $\eta_p^2 = .30$, with higher amplitudes corresponding to higher trait ratings. This is not in line with findings from Experiment 5.2, which showed that PC2 correlated with neuroticism only.

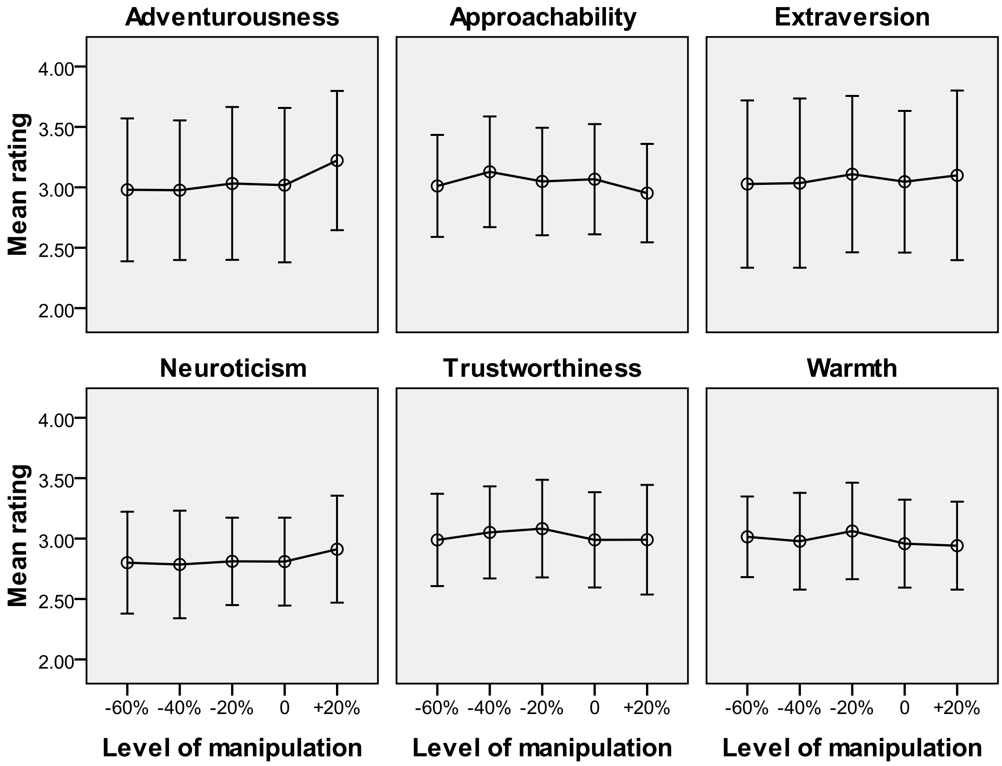


Figure 18. Mean trait ratings (± 1 *SD*) on six scales for five levels of manipulation of amplitude of PC2 scores.

General Discussion

The overall aim of the experiments reported in this chapter was to assess the link between bodily motion and perceived personality based on point-light walkers. In Experiment 5.1, we showed that point-light walker stimuli depicting single gait cycles were reliably rated on six personality traits. This finding adds to a range of studies showing consensus at zero acquaintance (Albright, et al., 1988; Albright, et al., 1997; Ambady & Rosenthal, 1993; Engell, et al., 2007; Heberlein, et al., 2004; Kenny, et al., 1994).

Results from Experiments 5.3.1 and 5.3.2, in which PC1 and PC2 were individually manipulated, partially confirmed the link between motion parameters and perceived personality. The systematic manipulation of the first component (PC1) had an impact on trait ratings, in line with the findings from Experiment 5.1. The impact of manipulation of PC2 was weak and, contrary to the manipulation of PC1, did not corroborate the findings from Study 1. It could be that the higher magnitude of manipulations required for PC2 made the walkers less natural-looking. This may not be surprising given that the second principal component, on average, accounted for 6.6% of the variance of the motion data. Our study did not simultaneously manipulate PC1 and PC2. Although it is possible that personality impressions depend on interactions of these two parameters, the weak correlation coefficients from Study 1 made us abandon such an elaborate design. However, future studies, with different stimuli, may allow investigating whether such interactions are possible.

In summary, we have shown that point-light walker files can indeed be modified in order to influence perceived personality traits of target walkers. This outcome may be of relevance to the creation of computer avatars used for entertainment or marketing. It may be of relevance also to human interactions, since first impressions are automatic (Asch, 1946; Bar, et al., 2006; Willis & Todorov, 2006) and can lead to important social decisions (Kramer, et al., 2010; Little, et al., 2007; Todorov, et al., 2005; van't Wout & Sanfey, 2008).

However, it is not certain that the minimal cues identified in this chapter can be ‘taught’ or, if they can, whether instructions to move in certain ways are effective; e.g., will perceived extraversion increase if you instruct the target to walk with arms extended? Further studies may be required to verify this. We also do not know whether people use bodily motion as cues for personality when information such as facial expressions, clothing or verbal behaviour is available to the observer. Findings have shown a kernel of truth for extraversion and conscientiousness using a zero acquaintance paradigm when the confederate is presented with audio recordings alone (Borkenau & Liebler, 1992). Validity increased as confederates were given access to richer stimuli, the highest being for video clips where the target could be heard reading a weather forecast. This is indeed indicative of cues other than body motion being of importance for personality trait ratings. However, studies have also shown that important social decisions can be traced down to movement cues (Kramer, et al., 2010). This means that gait information may have an impact on person perception and even social interaction also when other visual or verbal cues are available to an observer.

The experiments reported in this chapter have not explored the kernel of truth hypothesis (Berry, 1990; Bond Jr, Berry, & Omar, 1994; Penton-Voak, et al., 2006); that is, whether there is correspondence between how observers rate targets' personality traits and how targets rate themselves. Despite the lack of an association between motion parameters and self-reported personality, we investigate the kernel of truth hypothesis in Chapter VII because observers' trait impressions may rely on other cues not detected by our motion analyses. In Chapter VII we also investigate other variables potentially mediating the link between motion parameters and personality trait ratings.

Chapter VII: Investigating Non-Motion Cues for Personality Trait Judgements

In Chapter VI we replicated consensus at zero acquaintance for point-light walkers. This consensus led to the exploration of possible underlying cues for trait impressions. An initial computational motion analysis indicated that two parameters—the amplitudes of the first and second components—were linked with personality trait impressions. This link was further validated in an experimental design whereby the amplitudes were individually manipulated.

In the same chapter, we failed to find a link between self-reported personality and gait, by comparing the target individuals' scores on the NEO FFI and the motion parameters identified using PCA and sinusoidal fitting. However, we hesitate to conclude that there is no link between gait and personality, as it is possible that such a link exists but that our chosen motion analyses simply failed to find it (see Chapter IX for a discussion). In the current chapter, we therefore pursue the task of explaining the consensus further by asking if there is a link between the personality trait impressions and the actual personality traits of the targets. This is done by comparing the rating data from Experiment 5.1 (see Chapter VI) to the scores on the NEO FFI (Costa & McCrae, 1992) previously collected from the target walkers (see Chapter V).

Secondly, we investigate what underlying processing might be driving, or be associated with, the trait impressions. To this end, three further rating experiments were carried out (Experiments 6.2.1, 6.2.2 and 6.2.3). Across the three experiments, we assessed to what degree, if at all, personality trait ratings were associated with six predictor variables; that is, judgements of age, health, physical attractiveness, masculinity, arousal and valence.

Experiment 6.1: Correlations between Self-Reported and Peer-Reported Personality Traits—Investigating the Kernel of Truth Hypothesis

There are advantages in being able to accurately predict personality traits based on nonverbal behaviour (Barrett, Todd, Miller, & Blythe, 2005) and recent studies lend support to the idea that people are accurate in their ratings of certain traits based on nonverbal cues from photographs of faces (Penton-Voak, et al., 2006) or video clips (Borkenau, et al., 2004). To assess whether such accuracy could be obtained for whole-body motion, we correlated the self-reported personality scores of the targets with the observers' trait ratings. This approach is the most common for studies that assess the kernel of truth hypothesis (Borkenau & Liebler, 1992; Borkenau, et al., 2004; Penton-Voak, et al., 2006).

Even though no conclusive relationship was found between self-reported personality and the computer-extracted motion parameters in Chapter V, we decided to study the validity of trait judgements because of the possibility that a link between self-reported personality and gait exists but that our analyses simply failed to detect it. Firstly, the study had limited power due to the low number of walkers. Secondly, methods other than PCA exist for describing physical features of visual body stimuli: for example, in addition to subjective ratings, both posture cues, such as joint angles, and movement cues, such as a change in the linear weights of different body parts relative to neutral walking, have been found to classify emotional expressions (Roether, et al., 2009).

Method and Results. The Big Five scores from the NEO FFI SF (Costa & McCrae, 1992) administered to targets prior to the motion capture procedure were used as self-reported traits (see Chapter V). These were evaluated in relation to their corresponding rating scales from the observers, obtained in Experiment 5.1 (Chapter

VI). A judgement was deemed accurate if the trait rating of, say, neuroticism correlated with self-reported neuroticism. The data for male and female target individuals were analysed separately to avoid inflated correlations: gender differences in self-ratings of personality have indeed been found to coincide with social stereotypes (Costa, et al., 2001).

No validity of trait judgements was found, in that there were no significant within-trait correlations between self- and peer- ratings: coefficients ranged from −.17 (openness to experience) to .23 (conscientiousness). There were no self-other correlations for any other trait scale combinations once Bonferroni correction was applied. For female targets, all absolute ρs < .35, all *p*s > .23. However, for male targets only, self-reported extraversion was negatively correlated with both perceived trustworthiness ($\rho = -.74$, $p = .006$) and warmth ($\rho = -.60$, $p = .039$); apart from the extraversion construct, all other absolute ρs < .43, all *p*s > .17. For these analyses 30 correlations were performed (15 for each gender) and, according to the Bonferroni correction method, *p*-values should thus not be deemed significant unless they reach .0017; the interpretation of this coefficient should therefore be handled with caution.

Discussion. The findings contrast with studies reporting a kernel of truth in ratings of personality traits based on limited behaviour (Borkenau, et al., 2004; Penton-Voak, et al., 2006). We acknowledge that our limited sample size may not have given this particular analysis in our study enough power. However, most previous studies reporting a kernel of truth when a confederate makes personality judgements of a stranger have allowed for different cues to be available to the confederate, such as the face (Penton-Voak, et al., 2006) or even verbal information (Borkenau, et al., 2004). Indeed, Montepare and Zebrowitz-McArthur (1988) did not find support for a kernel of truth when assessing perceived personality of point-light walkers; likewise, Kenny, Horner, Kashy, and Chu (1992) found no validity in trait judgements when observers were shown 20-s silent video tapes of targets who were seated and unaware of being recorded.

It appears that trait impressions based on motion data of point-light walkers are reliable but not valid: observers agree with each other about which walkers look, say, extraverted or conscientious, but their impressions do not correspond to how the targets rated themselves. It could be that there is a kernel of truth in personality judgements only for other types of information (e.g., verbal or facial stimuli). Even with other types of stimuli, however, there is mixed evidence for the kernel of truth hypothesis, and the impact of transient cues on trait impression (Knutson, 1996; Montepare & Dobish, 2003) may indicate that this is highly context-dependent.

Not only was there no conclusive link between self-reported and perceived personality on the individual personality traits, but also the analyses did not find any link between any of the self-reported personality scores with any of the trait ratings. One exception was found for self-reported extraversion of the male targets. There is evidence that self-reported extraversion may be involved in aspects of body motion for males: in one study, self-reported extraversion using the NEO FFI was found to be positively linked to the perceived quality of someone's dance (Fink, et al., 2012). This does not, however, explain how, in the current experiment, the link between extraversion and the positive traits of trustworthiness and warmth was negative. This observation, coupled with the fact that the correlation did not survive a Bonferroni correction, leads us to conclude that there were no links between self-reported and perceived personality for any of the trait-combinations in our data.

However, caution must be shown when generalising these findings, and not only due to the limited sample size in the study: it is important to bear in mind that self-report questionnaires are not the only way to measure personality, nor is the NEO FFI SF the only questionnaire. The strongest opponents of self-reports as a source of personality descriptions claim that it is the judgements made by peers that are the true measurement of personality (Funder, 1995; Swann, 1984). Others argue that external criteria such as job performance should be used to assess personality (Oh, et al., 2011; Zimmerman, et al., 2010). Finally, the Self-other Knowledge Asymmetry model (Vazire, 2010) suggests that individuals know themselves best on some traits (e.g., neuroticism) whilst others are better judges on traits that may have observable criteria (e.g., intellect). We therefore cannot draw the conclusion from the lack of significant correlations between self- and peer-ratings in our study that there is no validity in trait judgements based on bodily stimuli. A follow-up study would indeed benefit from using ratings of target individuals by known peers, or other external criteria such as job performance, for assessing validity.

If we conclude that there is no link between self-reported personality of our targets and the perceived personality based on the point-light walker stimuli, yet we have consensus in personality trait judgements (see Experiment 5.1, Chapter VI), this means that observers base personality trait impressions on factors not related to the self-reported personality traits of the walkers. In Chapter VI we showed that motion parameters could be extracted from the point-light stimuli to predict, and influence, personality trait impressions. However, we have not explored what type of processing occurs within the observers when they attribute personality traits. We know that observers base trait judgements on factors such as emotions (see Chapter III), youthfulness (Montepare & Dobish, 2003) and physical attractiveness (Dion, et al., 1972). In order to assess whether any such factors mediated personality trait impressions in the present study, three further experiments were therefore carried out.

Experiments 6.2.1, 6.2.2 and 6.2.3: Additional Trait Ratings of Point-Light Walkers

The results of Experiment 6.1 showed that the self-reported personality traits of the walkers were not the driving factors for trait impressions, either directly or indirectly. This finding is further strengthened by the results from Experiment 5.2 (Chapter VI), which showed no relationship between self-reported personality and any of the extracted motion parameters. In three experiments (Experiments 6.2.1, 6.2.2 and 6.2.3) we therefore investigated whether alternative decision making may mediate the trait impressions of point-light walkers. To do this we assessed to which degree, if at all, impressions of personality were associated with judgements of age, health, physical attractiveness, masculinity, arousal and valence (henceforth *predictor variables*).

The chosen predictor variables have been shown to be associated with personality and first impressions to some extent. We collected ratings of age because this has been found to be associated with perceived personality in point-light walkers (Montepare & Zebrowitz-McArthur, 1988). Attractiveness has been found to predict personality impressions, such as extraversion ratings based on faces (Albright, et al., 1988). Attractiveness is closely linked to perception of health, consistent with the 'good genes' hypothesis, which posits that healthy individuals are attractive because of the importance for mate selection (B. C. Jones, et al., 2001). Health ratings based on body motion have also been found to predict voting choice (Kramer, et al., 2010). We therefore collected data on the perceived health of the walkers. Masculinity ratings were also collected, as personality ratings may be due to different social stereotypes for males and females; for instance, males are typically seen as less neurotic and more extraverted than women (Williams, et al., 1999).

Finally, we measured perceived arousal and valence to obtain measures of emotion. Indeed, it has been found that people attribute emotion to emotionally neutral stimuli and this can fuel subsequent trait judgements of faces (Montepare & Dobish, 2003). In Chapter III we confirmed that this also occurs for body-motion. Point-light walkers have further been shown to be more easily recognised as humans when the targets were instructed to appear angry (Chouchourelou, et al., 2006), confirming the interconnection of neurological structures involved in analysis of human motion and of emotion (e.g., Heberlein, et al., 2004).

Method. In Experiment 6.2.1, we collected data on the perceived masculinity of each walker. Observers (n = 15; 3 males and 12 females; Mean age = 19.0 years, SD = 0.9) were shown the stimuli in a random order and asked to indicate which gender they believed the walker to be by using the keyboard, thus using the same procedure as Montepare and Zebrowitz-McArthur (1988). For each stimulus, masculinity was calculated as the proportion of participants who judged a stimulus to be ‘male’.

In Experiment 6.2.2, a different group of participants (n = 21; 3 males and 18 females; mean age = 19.7 years, SD = 6.0) was asked to estimate the targets’ age. Participants first undertook six practice trials in order for them to form an impression of the mean age of the targets. These practice clips had been previously selected to vary in perceived age on the basis of the judgements of four separate observers. The practice stimuli were sampled from the same target walks used in the experimental trials, but different versions of the walks were used during the practice. Following the practice trials, participants were shown all 26 stimuli randomly, and asked to indicate how old they perceived the walker to be by writing their answer on a piece of paper. Exact perceived age, rather than a rating scale, was used as the dependent variable in order to follow the procedures of Montepare and Zebrowitz-McArthur (1988), who found youthfulness to predict personality trait impressions. The mean judged age across all participants was used as our final measurement of youthfulness.

Finally, in Experiment 6.2.3, a third group of participants (n = 22; four males and 18 females; mean age = 19.7 years, SD = 4.8) rated the target walkers on physical attractiveness, health, arousal and valence on five-point rating scales. The antonym-pairs ‘unattractive-attractive’, ‘unhealthy-healthy’, ‘calm-excited’ and ‘unpleasant-pleasant’ were used.

All further methods, as well as apparatuses, were identical to those employed in Experiment 5.1, Chapter VI.

Results and Discussion. Inter-rater reliability was high for all predictor variables: masculinity (α = .86), age (α = .89), attractiveness (α = .86), health (α = .94), arousal (α = .95) and valence (α = .87).

Accuracy was above chance for judgements of gender of both male (one-sample t-test: t_{10} = 9.15, p < .001) and female (t_{12} = −3.34, p < .01) walkers. Male walkers were correctly classified in 86.0% of the trials whereas females were correctly classified in 70.4% of the trials. The higher classification accuracy for males than for females fits with a male bias reported in the literature (Troje & Szabo, 2006). Perceived age of the targets (M = 35.4, SD = 8.5) varied highly between targets (range: 23.1 – 56.8) and was significantly higher than their actual age (M = 19.3, SD = 0.9), Z = 4.5, p < .001, r = .87. The two variables were not correlated: (ρ = −.01, p = .97). It is possible that, due to the sampling procedure the age range of the target walkers may not have been sufficient to detect any such accuracy (age demographics of target walkers can be seen in Chapter V).

The two motion parameters that were found to drive personality ratings in Chapter VI (i.e., amplitudes of PC1 and PC2) showed a pattern of correlations with the predictor variables. A dissociation was found whereby PC1 amplitude correlated with attractiveness, health, valence and perceived age, whilst PC2 amplitude correlated with arousal and masculinity (see Table 28).

Table 28

Correlations (ρ) between Component Amplitudes and Predictor Variables

	PC	
Rating scale	1	2
Masculinity	<.01	$.57^{**}$
Age	$-.40^{*}$	.38
Attractiveness	$.46^{*}$	−.23
Health	$.48^{*}$	−.36
Arousal	.23	$-.53^{**}$
Valence	$.58^{**}$	−.18

Note. PC = principal component number. * $p < .05$; ** $p < .01$. Columns represent correlations between trait ratings and the amplitudes of scores the respective component as determined through sinusoidal fitting.

Strong correlations were also seen with the previously collected trait ratings. For instance, extraversion correlated with all scales (all absolute ρs > .57, all *p*s <.003) apart from masculinity ($\rho < .01$, $p = .98$). See Table 29.

Table 29

Correlations (ρ) between Predictor Variables and Personality Trait Ratings

Rating scale	Masc.	Age	Attr.	Heal.	Arou.	Val.
Age	.23					
Attractiveness	−.13	$-.83^{***}$				
Health	−.33	$-.85^{***}$	$.86^{***}$			
Arousal	$-.40^{*}$	$-.67^{***}$	$.60^{**}$	$.77^{***}$		
Valence	−.16	$-.81^{***}$	$.71^{***}$	$.80^{***}$	$.59^{**}$	
Approachability	.12	−.23	$.44^{*}$	.16	−.27	.32
Extraversion	.00	$-.69^{***}$	$.57^{**}$	$.67^{***}$	$.58^{**}$	$.82^{***}$
Neuroticism	$-.57^{**}$	−.39	.30	$.51^{**}$	$.76^{***}$	.25
Novelty−preference	−.23	$-.70^{***}$	$.55^{**}$	$.73^{***}$	$.74^{***}$	$.79^{***}$
Trustworthiness	−.15	$-.53^{**}$	$.60^{**}$	$.54^{**}$	.22	$.68^{***}$
Warmth	−.13	$-.46^{*}$	$.49^{*}$	.38	.14	$.68^{***}$

Note: Masc. = Masculinity; * $p < .05$; ** $p < .01$; *** $p < .001$.

Stepwise regression analyses were carried out for each personality trait individually (see Table 30). The results revealed that the variables chosen to obtain an estimate of perceived emotion (i.e. arousal and valence) were retained as predictors of all personality traits. For instance, valence was kept as a predictor in all personality trait ratings apart from approachability. High arousal was associated with perceived novelty-preference and neuroticism, whilst low arousal was associated with perceived approachability and warmth. Two other rating variables unrelated to emotion were also kept by the analyses: masculinity predicted perceived emotional stability (neuroticism complemented), whilst attractiveness predicted approachability. The final two predictor variables, age and health, were not retained by any of the regression analyses.

Table 30

Stepwise Regression Analyses with Predictor Variables and Personality Ratings

	Predictor variable				
Pers. trait	Arousal	Valence	Masculinity	Attract.	R^2
Approch.	–0.81 (5.85***)	--	--	1.06 (7.65***)	0.73
Extrav.	--	0.79 (6.33***)	--	--	0.61
Neuro.	0.89 (4.97***)	–0.49 (2.99**)	–0.31 (2.39***)	--	0.70
NP.	0.50 (4.23***)	0.50 (4.22***)	--	--	0.84
Trust.	--	0.76 (5.65***)	--	--	0.57
Warmth	–0.38 (2.13*)	1.00 (5.58***)	--	--	0.62

Note. Standardised β-coefficients (*t*-values in brackets) for six separate stepwise regression analyses. Six predictor variables were entered into the stepwise regression (age and health were not retained by any of the analyses). *R*2 = goodness of fit for the given model. * $p < .05$; ** $p < .01$; *** $p < .001$. Pers. = personality, NP. = novelty-preference, Approch. = approachability, Extrav. = extraversion, Trust. = trustworthiness, Attract. = attractiveness.

It thus appears that personality ratings may have been mediated by impressions of emotion, reflected in the number of stepwise regression analyses that kept the predictors arousal and valence (either alone or combined). Judgements of masculinity and attractiveness were found to predict emotional stability and approachability, respectively. We note that a high degree of colinearity was found, with many of the predictor variables strongly correlated; for instance, as can be seen in Table 29, attractiveness and health showed a correlation coefficient of .86 ($p <$ .001). This multicolinearity—which must be expected for rating scale variables of this type—means generalisations drawn from the regression analyses should be treated with caution. However, overall, the outcome from the multiple regressions is consistent with previous research on person perception and fits with empirical evidence that perception of emotion from emotionally neutral stimuli can affect personality attribution (Montepare & Dobish, 2003).

General Discussion

In Chapter VI, we showed that people make reliable personality trait judgements on point-light walkers (Experiment 5.1) and we identified motion parameters that predicted these trait judgements (Experiment 5.2). In the current chapter, we set out to explore what other underlying factors were driving the agreement, by testing both psychological factors intrinsic to the target walkers (Experiment 6.1) and by assessing other judgements involved in the processing of personality traits by the observers.

We found no link between the personality traits of the walkers themselves, as assessed through the NEO FFI SF (Costa & McCrae, 1992), and the perceived personality traits, as assessed through our trait ratings. However, we found a number of predictor variables, indicative of other judgements driving personality trait impressions. Consistent with previous research, perception of emotion, masculinity and attractiveness was tied with personality trait judgements.

In Experiment 5.2 (Chapter VI) we identified two motion parameters linked with the personality trait ratings. Experiments 5.3.1 and 5.3.2 (Chapter VI) showed that this association was not simply a by-product of some other underlying variable (e.g., a static feature that both influences motion parameter and drives trait impressions) but that the motion parameters could, in fact, explain and influence perceived personality traits. If the retained predictor variables are interpreted as mediating variables, one should expect strong correlations between these variables and the motion parameters. However, comparing the analyses from Experiments 6.2.1, 6.2.2 and 6.2.3 with findings from Experiment 5.2 (Chapter VI), one can see that perceived emotion, masculinity and attractiveness correlated much less strongly with PC1 and PC2 amplitude than did the corresponding personality trait ratings. This may be indicative of other motion parameters being more important for conveying emotion; for instance dynamic asymmetry (Roether, et al., 2008).

As stated above, we based the classification of predictor and predicted variables on previous research. For instance, we have reason to believe that it is perception of emotion that influences perceived personality, and not the other way around. However, with high degrees of colinearity, such as those reported here, caution must be shown when interpreting outputs from regression analyses. It could be that the emotion judgements are intertwined with the personality judgements, rather than driving them.

The issue of multicolinearity is dealt with in Chapter VIII, where we perform a PCA on all personality trait ratings *and* the predictor variables. This chapter also serves the purpose of investigating whether a similar pattern of personality trait attribution occurs for the stimuli used in Part 2 (i.e. the point-light walkers) of the thesis as those used in Part 1 (i.e. the patch-light and full-light clips in which the targets were instructed to display an emotion or an action).

Chapter VIII: Pattern of Trait Ratings on Point-Light Walkers

In the previous chapters we presented data of perceived personality (Chapter VI) as well as ratings on other traits to collect predictor variables (Chapter VII) based on point-light walker stimuli. Chapter VII showed how impressions of emotion—as measured through arousal and valence—as well as physical attractiveness and masculinity, predicted personality trait attribution. However, high degrees of colinearity were found between the variables, which meant that any conclusions drawn from the regression analyses had to be treated with caution. The current chapter serves two purposes. First, we further the research question from Chapter VII, which was to find other trait impressions that may be underlying personality trait judgements. A new analysis is introduced in order to deal with the issue of multicolinearity. Second, we assess whether the pattern of trait ratings on point-light walkers corroborates those presented in Chapter IV for full-light and patch-light stimuli.

Experiment 7.1: Principal Component Analysis on All Rating Data for Point-Light Walkers

The literature on the dimensionality of personality perception largely agrees that observers cluster personality judgements along two dimensions. The first is typically labelled valence, proposed to reflect the observer gauging the intention of an actor to cause harm; the second is typically labelled dominance, proposed to reflect the observer assessing the ability of this actor to cause harm (Asch, 1946; Fiske, et al., 2007; Oosterhof & Todorov, 2008; Rosenberg, et al., 1968; Todorov, et al., 2008). Research on emotion recognition provides preliminary evidence that early assessment of intention is beneficial for survival**;** faces are considered to convey intentions through expressions of emotion (for a review see, e.g., Parkinson, 2005) and expressions of anger, happiness and surprise—which may be considered indicators of someone's immediate intentions—contain diagnostic features that in principle are available at greater distances (Smith & Schyns, 2009). It has also been suggested that male facial expressions of anger are accurately transmitted from over 100 metres distance, farther than sadness or fear (Hager & Ekman, 1979; although, see Smith & Schyns, 2009).

Since observers have access to dynamic body information derived from someone's walking style at greater distances than to static information from faces—e.g., in times of reduced visibility—it is reasonable to assume that a similar processing occurs for gait; that is, that impressions relating to intention (valence) will precede impressions relating to ability (dominance). It has indeed been shown that people are more sensitive to anger in point-light walkers than to other emotions (Chouchourelou, et al., 2006).

However, it is not certain that the gait obtained through our recording procedures contains enough information for observers to gauge intention. In order to test the hypothesis that judgements of intention precede judgements of ability when observers are shown emotionally neutral point-light walkers, and to see whether the findings from Chapter IV could be corroborated, we submitted the rating data of all experiments presented in Chapter VI and VII to a PCA to explore underlying, latent variables.

Method. Due to the extra analyses performed on point-light walkers in Chapter VII, we had more rating data than was available for the PCA on full-light and patch-light stimuli reported in Chapter IV. Twelve variables were entered into the PCA. Six of these were the personality trait ratings obtained in Experiment 5.2: approachability, extraversion, neuroticism, novelty-preference, trustworthiness and warmth (see Chapter VI). The other variables were the predictor variables identified in Experiments 6.2.1, 6.2.2 and 6.2.3: perceived age, arousal, attractiveness, health, masculinity and valence (see Chapter VII). As was the case for the PCA presented in Chapter IV, Varimax was used as rotation method.

Results and Discussion. A PCA including all 12 variables yielded a three-factor solution, accounting for 88.2% of the variance in the data. The first component accounted for 58% of the variance whilst the second accounted for 21.2% of the variance in the rating data.

The PCA split up the personality trait scales so that extraversion and novelty-preference formed one group (PC1) whilst approachability, trustworthiness and warmth formed another (PC2). This grouping fits with the findings reported in Chapter IV, where a PCA was performed on all personality ratings made on full-light and patch-light displays. Two differences emerged between the PCA reported in Chapter IV and the one reported here: the order of the PCs was reversed and secondly, emotional stability (neuroticism complemented) was now separated from the other traits, forming PC3. See Table 31.

Table 31

Factor Loadings for PCA Performed on all 12 Rating Scales

	Principal component		
Rating scale	1	2	3
1) Personality traits:			
Approachability	−.10	**.92**	−.21
Extraversion	**.91**	.21	−.13
Neuroticism	.32	−.32	**.84**
Novelty-preference	**.94**	.12	.17
Trustworthiness	**.44**	**.80**	.06
Warmth	.39	**.83**	−.03
2) Predictor variables:			
Age	**−.72**	**−.53**	−.22
Arousal	**.84**	−.01	**.47**
Attractiveness	**.60**	**.68**	.16
Health	**.83**	.36	.35
Masculinity	−.04	−.13	**−.91**
Valence	**.80**	**.51**	.11

Note. Boldface indicates factor loadings stronger than .40

Perceived age, attractiveness and valence had strong loadings on both PC1 and PC2. For this reason another PCA was calculated excluding these variables. This produced a three-factor solution with the same pattern, accounting for 89.7% of the variance in the data. The first component accounted for 50.5% and the second 27.3% of the data. Subsequent PCAs are based on this latter analysis.

PC1 had high loadings of arousal and health (see Table 32). This fits with the idea that this judgement is related to perceived dominance because research on face perception suggests that perceived dominance reflects the perceived ability to cause harm (Oosterhof & Todorov, 2008; Todorov, et al., 2008).

Table 32

Factor Loadings for PCA Performed on Nine Rating Scales

	Principal component		
Rating scale	1	2	3
1) Personality traits:			
Approachability	−.11	**.90**	−.18
Extraversion	**.91**	.24	−.13
Neuroticism	.33	−.36	**.82**
Novelty-preference	**.95**	.13	.17
Trustworthiness	.40	**.82**	.10
Warmth	.39	**.86**	.01
2) Predictor variables:			
Arousal	**.85**	−.04	.42
Health	**.82**	.33	.34
Masculinity	−.05	−.11	−**.92**

Note. Boldface indicates factor loadings stronger than .40

Factor scores were calculated in order to describe the components using the findings from the effort-shape analysis reported in Experiment 4.2 (Chapter V). The first component correlated strongly with the effort-shape variables energy, space, limb and torso (see Table 33), thus representing a *channelled*, *powerful* walk with high use of personal space. Once corrected for multiple comparisons, this factor score correlated with the motion parameter PC1 amplitude only.

Table 33

Correlations (ρ) between Rating Factor Scores and 1) Motion Parameters and 2) Effort-Shape Parameters

	Principal Component		
Rating scale	1	2	3
1) Motion parameters:			
Motion-PC1 Amplitude	.47*	.40*	−.04
Motion-PC1 Frequency	.21	−.09	.52**
Motion-PC2 Amplitude	−.11	.18	−.61**
Motion-PC2 Frequency	.46*	.02	.44*
2) Effort-shape parameters:			
Torso	.91***	.05	.01
Limb	.79***	−.10	−.09
Energy	.72***	−.39*	.08
Space	.64***	−.16	.55**
Time	.50*	−.12	.69***
Flow	.03	−.39	.75***

Note. Spearman's correlation coefficients between factor scores obtained using regression method and Varimax rotation, from PCA on nine rating scales and 1) motion parameters and 2) effort-shape parameters (reported in Chapter V). * $p < .05$; ** $p < .01$; *** $p < .001$

PC2, formed by approachability, trustworthiness and warmth, did not see exclusive loadings by any of the predictor variables. The grouping of these traits corroborates the findings from Chapter IV, and coincides with research on face perception, closely resembling a valence component (Oosterhof & Todorov, 2008;

Todorov, et al., 2008). Once corrected for multiple comparisons, the factor score did not correlate significantly with any of the effort-shape or motion parameters (see Table 33). The factor score correlated with attractiveness (ρ = .42, p = .34) and valence (ρ = .42, p = .03); however, these coefficients were weak and did not survive a Bonferroni correction for multiple comparisons.

PC3 saw a negative loading of masculinity (−.92), which was stronger than the loading of Neuroticism (.82; see Table 32). The differential loading strength suggests that observers first made a judgement of how masculine they found the target before making up their mind about how anxious this person was, consistent with a social stereotype of women being seen as more neurotic than men (Williams, et al., 1999). We therefore tentatively labelled the third component a gender construct. The factor score showed dissociation with the first component in terms of Effort-Shape parameters: the PC3 factor score correlated with flow and time, which were both unrelated to the other factor scores. *Slow* and *relaxed* walks were thus seen as more *masculine* and less *neurotic*. The gender factor score also correlated with the effort-shape parameter space, signifying that *direct*, *focused*, walks were associated with higher ratings of this component, and thus a more masculine gait.

To see whether the factor scores could be approximated by linear combinations we calculated the arithmetic mean of the variables loading on the respective components. A dominance construct was defined as the mean of extraversion, novelty-preference, arousal and health. A valence construct was defined as the mean of approachability, trustworthiness and warmth. Finally, a gender construct was defined as the mean of masculinity and emotional stability (neuroticism complemented). There were strong correlations between the factor scores and their approximated constructs (all ρs > .90, all ps < .001). The gender construct saw a weak, marginally significant, correlation with dominance ($\rho = -.40$, $p = .04$); the constructs were otherwise not correlated (ρs < .37, ps > .07). These weak correlations (or trends) are consistent with previous findings from studies of trait ratings from face stimuli (e.g., Oosterhof & Todorov, 2008).

Recall that the predictor variable masculinity was calculated by the ratio of observers classifying the given target walker as male and observers were generally accurate in this perception (see Chapter VII). This led us to perform further PCAs for male and female walkers separately. Eleven variables were now used: all six personality trait scales as well as all the predictor variables bar one: masculinity. However, very few variables were retained in the rotated component solutions, probably due to small sample sizes (14 female and 12 male target walkers; see Chapter V). PCAs were carried out separately for males and females. If a variable had a loading of more than .4 on more than one component, a new PCA was carried out without this variable. The final two-component solutions were similar for both genders in that none of the predictor variables was kept; however, they differed with respect to which personality trait ratings were kept (see Table 34 and Table 35).

Table 34

Variance Explained by Components for PCA Solutions by Male and Female Walkers

	Male walkers		Female walkers	
PC	Perc. var.	Cum. var.	Perc. var.	Cum. var.
1	49.3%	49.3%	57.7%	57.7%
2	33.6%	82.9%	35.1%	92.8%

Note: Only principal components with eigenvalue > 1 are shown. PC = principal component number, Perc. var. = percentage of variance explained, Cum. var = cumulative percentage of variance explained.

Table 35

Factor Loadings for PCAs Carried out on Male and Female Walkers Separately

	Male walkers		Female walkers	
Rating scale	PC1	PC2	PC1	PC2
Approachability	.05	**.83**	−.11	**.94**
Extraversion	**.95**	.20	**.98**	.14
Neuroticism	−.07	**−.85**		
Novelty-preference	**.97**	−.06	**.98**	.03
Warmth			.39	**.85**

Note. PC = principal component. Boldface indicates factor loadings stronger than .40

Despite the discrepancies between the PCAs carried out for male and female target walkers in terms of the loading pattern of the personality traits, it appeared that, for both genders, observers made two judgements: 1) a judgement resembling the dominance construct: extraversion and novelty-preference and 2) a judgement resembling the valence construct: approachability, emotional stability (for male walkers) and warmth (for female walkers). This trend can thus be said to support the findings from the overall PCA carried out across male and female walkers. However, trustworthiness was not kept by the analyses because it did not have a clear loading pattern. These discrepancies may be due to a limited number of stimuli within the samples.

General Discussion

This chapter presented a PCA performed on all trait ratings collected on point-light walkers in Chapters VI and VII. This was done to deal with the issue of colinearity of predictor and predicted variables in Chapter VI, and to corroborate findings from Chapter IV; more specifically, to assess whether patterns in personality trait judgements for point-light walker stimuli were similar to the pattern for patch-light and point-light stimuli presented in Part 1 of the thesis.

In summary, the findings from Experiment 7.1 suggest that the main personality judgement formed by observers when viewing point-light walkers was an impression of extraversion and novelty-preference, which was partly driven by impressions of arousal and health. Based on previous literature (Fiske, et al., 2007; Oosterhof & Todorov, 2008; Todorov, et al., 2008) we believe this component reflects a dominance construct. Effort-Shape analyses indicated that this impression may be affected by high use of personal space.

The second judgement regarded the personality traits approachability, trustworthiness and warmth, and it is possible that these traits represent a valence construct. This was further supported by a correlation, albeit weak, with attractiveness and valence. The fact that the correlation with valence was only weak (and non-significant after correction for multiple comparisons) may seem puzzling. However, it could be that the chosen antonyms (unpleasant-pleasant) were not the most appropriate for tapping into this construct. Indeed, Oosterhof and Todorov (2008) suggest that *trustworthiness* is the trait scale that best encompasses a judgement of valence. We found no motion parameter associated with this component; neither from subjectively collected variables nor from computer-extracted parameters (see Chapter V). It is likely that other motion cues not detected by our analyses drive the valence judgement; further analyses would be needed to assess this.

Emotional stability (neuroticism complemented) formed its own component and analyses suggest that a judgement of masculinity may be associated with this personality trait. We believe the perception of target gender may have driven the consensus in ratings of neuroticism: high masculinity was associated with perception of emotional stability. This fits with studies on gender stereotypes (Williams, et al., 1999). This was confirmed through further analyses: Performing PCAs on male and female targets separately removed the third component. It appears that emotional stability is an impression intrinsic to perceived valence for male walkers only. Note, however, that arousal and valence remained across all PCAs, and the masculinity component accounted for only 12% of the variance in the overall PCA.

The findings therefore confirm to some extent those from Chapter VII (Experiments 6.2.1, 6.2.2 and 6.2.3), but they also show how the choice of analysis of data may change the interpretation. The findings further partly corroborate the results from Chapter IV, but an interesting discrepancy was found. That is, although the grouping was largely consistent, the order in magnitude, i.e. the percentage of variance explained, of the constructs was different. Whilst for full-light and patch-light displays a valence component was shown to explain the largest proportion of the data, in the point-light walker displays it was the dominance component that explained the most of the variance: more than twice that of valence. Furthermore, the component order found with the point-light walkers was opposite to that found in previous studies on the dimensionality of personality perception (Asch, 1946; Fiske, et al., 2007; Oosterhof & Todorov, 2008; Rosenberg, et al., 1968; Todorov, et al., 2008) including perception of faces (Oosterhof & Todorov, 2008; Todorov, et al., 2008). If observers are more likely to extract impressions of dominance rather than valence from gait, this suggests that walking is perceived to contain more information about power, as per the theory that dominance reflects the perceived ability to cause harm (Oosterhof & Todorov, 2008; Todorov, et al., 2008). This is supported by the strong link between perceived power and the PC1 (dominance) factor score in our data. The literature also suggests that gait effectively affects perceived vulnerability (Gunns, et al., 2002; Johnston, et al., 2004; Sakaguchi & Hasegawa, 2006), even if the perceived vulnerability is unrelated to the actual power of the walker (Gunns, et al., 2002; Johnston, et al., 2004).

The magnitude of the valence component was lower than expected and it is likely that this is due to gait containing fewer cues of intention and motivation. The second component, sometimes labelled ‘competence’ (see, e.g., Fiske, et al., 2007), is perceived to be a dispositional trait: c.f., ‘you cannot change your competence.’ Considering how people’s gait contains individual difference enough to correctly classify gender (Kozlowski & Cutting, 1977; Pollick, Kay, Heim, & Stringer, 2005; Troje, 2002) and recognise individuals (Cutting & Kozlowski, 1977; Troje, et al., 2005), the reverse ordering of the valence and dominance components may not be so surprising after all.

The discrepancy between the rating pattern for point-light walkers and for the whole-body stimuli presented in Chapter IV (viz. patch-light and full-light clips of targets displaying an emotion or an action) is still puzzling and warrants further investigation. A first step to explore possible explanations is to look at the differences between the stimuli in terms of physical aspects. As far as the observer is concerned, the clips used in Part 1 of the thesis were different from those used in Part 2 chiefly with respect to two aspects: 1) the displayed behaviour of the targets and 2) the display format employed for the final stimuli. Due to this double-manipulation it is hard to immediately pin-point the difference in the rating patterns found in Chapter IV and those reported in this chapter. Furthermore, sub-analyses in Chapter IV showed that similar patterns in trait attribution occurred across type of behaviour (action vs. emotion) and display format (patch-light vs. full-light).

Since patch- and full-light displays contain more static information than do point-light displays, further studies would need to be carried out to investigate trait attribution patterns for point-light stimuli of targets engaging in different actions or portraying given emotions. If the resultant patterns of findings were to resemble those obtained in Part 1 of the thesis, i.e. if the valence component were to account for most of the variance, it is likely that it was the instruction that produced the differences in trait attribution. If, on the other hand, the patterns of findings were to be similar to those found with the point-light walkers, there would be grounds to conclude that the difference in trait attribution was due to more static information being available to observers for patch-light and full-light compared to point-light stimuli.

Regardless of which interpretation is correct, the findings of the current chapter, based on experiments presented in Chapters VI and VII, suggest that the dynamic information extracted from our point-light walker stimuli is likely to cue first impressions of a person’s ability to cause harm and, second, their intentions of doing so. Due to the robust finding that first impressions typically involve first an assessment of intention and then an assessment of ability (see e.g., Fiske, et al., 2007), we believe that the rating pattern found for our point-light walkers is reflective of a lack of motivational cues in our point-light walker stimuli, rather than a difference in how observers process these stimuli. Finally, we stress that this conclusion cannot be generalised to gait in general. For instance, it is possible that giving the target walkers different instructions during the motion capture procedure (e.g., ‘walk angrily’) may have yielded different rating patterns.

DISCUSSION

Chapter IX: General Discussion

Background and Goals

It has been shown that people go beyond the information immediately available to them when making judgements about others (Allport & Postman, 1947; Uleman, et al., 2005), inferring traits from impressions of emotions (Montepare & Dobish, 2003), behaviours (Winter & Uleman, 1984), or static visual cues (Polinko & Popovich, 2001). Such initial impressions are made automatically (Bar, et al., 2006; Willis & Todorov, 2006) and can lead to important social decisions (Little, et al., 2007; Todorov, et al., 2005; van't Wout & Sanfey, 2008). Not only do we know that such judgements are made automatically, but it has also been shown that observers largely agree with one another when attributing personality to strangers based on nonverbal behaviour (Albright, et al., 1988; Borkenau & Liebler, 1992; Gunns, et al., 2002; Kenny, et al., 1992).

People readily assign personality-like traits to moving objects (Heider & Simmel, 1944). Importantly, dynamic cues have further been shown to influence personality trait impressions of humans (Montepare & Zebrowitz-McArthur, 1988) and social decisions such as voting choice (Kramer, et al., 2010). This means that body-motion cues can lead to both explicit and implicit personality judgements, a finding of relevance to anyone interested in person perception, including animators who seek to create believable characters.

Despite compelling evidence that body motion influences personality trait judgements, most studies investigating the contribution of body-motion cues to person perception typically identify cues to determine the sex (Kozlowski & Cutting, 1977; Pollick, et al., 2005) or the identity (Cutting & Kozlowski, 1977; Loula, et al.,

2005) of a target, or to classify or quantify the expressed emotion (Atkinson, et al., 2004; Dittrich, et al., 1996).

The aim of this thesis was to investigate the link between body-motion and personality trait impressions, based on short, silent movie clips depicting human targets. We investigated behavioural cues, such as expressed action or intended emotion as well as mediating factors for personality trait attribution, such as perceived emotional content, attractiveness or masculinity.

The reported studies used tightly controlled stimuli: silent movie clips of targets engaging in actions, expressing an emotion, or simply walking. Static information available to observers was strongly reduced: faces were obscured, and body form was minimised, or even removed, through the use of patch-light (see Chapter II) or point-light (see Chapter V) displays. Using these stimuli, we further aimed to replicate consensus at zero acquaintance (Albright, et al., 1988) and to explore the kernel of truth hypothesis (Penton-Voak, et al., 2006). Through both two-dimensional analysis—in which the use of personal space and general space was calculated through tracking the silhouette images of target walkers—and three-dimensional analysis—in which trajectories of Cartesian positions of body parts was modelled using Principal Component Analysis—we aimed to identify motion cues that drive personality trait impressions from non-verbal behaviour. Finally, we wanted to assess whether manipulating motion cues could change perceived personality traits. To our knowledge, the studies reported here are the first to employ computerised analyses of motion parameters as means to explain and manipulate perceived personality traits.

Findings

Personality trait judgements and cues of consensus. Inter-rater agreement was strong for all trait-rating experiments reported throughout the thesis, consistent with many previous studies that have reported consensus at zero acquaintance (Albright, et al., 1988; Borkenau & Liebler, 1992; Kenny, et al., 1992; Penton-Voak, et al., 2006). Our studies showed that observers agree also when they are exposed to dynamic cues absent of information from the face, clothing, or other static cues, such as body shape.

In Chapter II, we showed that the type of action displayed by a target had an impact on the personality trait attribution of that target. This finding adds to previous literature which indicated that observers judge people out of context (E. E. Jones & Harris, 1967; Uleman & Moskowitz, 1994); and is evidence that people make inferences based on simple behaviours when forming impressions of others.

Findings from Chapter III showed that, when a target was asked to display an emotion, the type of emotion had an impact on personality trait impressions. Moreover, ratings of emotional content from walk-cycles (Chapter VII) and action-clips (Chapter II), in which targets were *not* instructed to display emotion, also had an impact on personality impressions. These findings add to the literature showing an overgeneralisation effect whereby impressions of emotion affect subsequent personality trait impressions (Montepare & Dobish, 2003; Said, et al., 2009).

Motion analyses. Inter-rater agreement was high throughout the experiments reported in the thesis. This held when targets engaged in actions (Chapter II) and tried to express emotions (Chapter III) but agreement was also high when the individual behaviours (e.g., the different types of actions) were analysed separately. This means that observers used visual information available to them to drive their trait attributions on the targets. Because static information was limited, we propose that observers used motion for their personality trait impressions and an important aspect of the thesis was to link trait attribution to dynamic cues.

Part 1 of the thesis used full-light and patch-light stimuli, created from video footage obtained from a single camera. Motion analyses were performed directly on the final stimuli shown to observers. The analyses showed that parameters of motion within the general and personal space could predict personality trait impressions. Different behaviours manifest different motion (e.g., sad clips will show less general movements than happy clips), yet analyses showed that the identified motion parameters were valid also within the individual actions and emotions intended by the target.

Part 2 of the thesis used stimuli that captured targets' natural walk. This uniform behaviour was chosen because it is frequently observed by an audience; for instance, a job interview panel will rarely ask the applicant to perform a star jump, or attempt to look angry. It is furthermore not emotionally laden and does not, in itself, imply intentions. This was important to control due to the already discussed impacts of type of action or emotional display on personality trait attribution.

The display format of the stimuli were point-light displays (Johansson, 1973), which contain a lot less static body-form information than full-light or even patch-light displays. These point-light displays were obtained from three-dimensional motion data of 13 markers demarcating points on the target's body, and this richness of data allowed for a more elaborate motion analysis. Due to the success of this technique in classifying gender differences in gait (Troje, 2002) we opted for principal component analysis, which resulted in four scalable motion parameters. These motion parameters were also found to predict, and explain, personality trait impressions.

Using these scalable parameters, we showed that point-light walkers can be modified in a way to affect personality trait ratings. Our effort-shape analyses of the stimuli further suggested that increased use of personal space, through more expansive use of limbs, affected perceived extraversion and novelty-preference. This finding is interpreted to reflect an observer's judgement of a target's ability to cause harm, rather than the intention of doing so.

Patterns in personality trait attribution. Chapters IV and VIII showed that patterns in explicit personality trait attribution based on whole-body movement largely fit with trait attribution based on other limited stimuli, such as faces, and human trait attribution in general. That is, trait judgements may be narrowed down to two underlying constructs: valence and dominance (Fiske, et al., 2007; Oosterhof & Todorov, 2008; Todorov, et al., 2008). The first is thought to reflect someone's intentions, and the other reflects someone's ability to cause harm. An interesting difference in trait attribution patterns emerged between the stimuli used in Part 1 and those used in Part 2. Specifically, the point-light walkers (Part 2) cued firstly impressions of dominance and secondly valence, contrary to the patterns found across all the other display categories from Part 1 and to the literature on trait impressions in general. Rather than reflecting a tendency of observers to ignore cues of intention (or valence) from gait, we propose that the discrepancy in trait attribution pattern is due to a lack of intention-cues in our particular stimuli.

Kernel of truth in personality trait attribution. As shown in Chapter VI, no accuracy was found in trait ratings of point-light walkers, in that there was no link between observers' trait impressions and target individuals' scores on a standardized and widely-used personality questionnaire. Chapter V further found no link between self-reported personality scores and motion parameters extracted through PCA. However, we tread with caution before generalising our findings to claim that body motion is unrelated to personality.

Other movements may contain more individual differences than simple walking, such as when someone engages with a target (e.g., picking up a box). Having a walker start from standstill is likely to contain other cues such as acceleration, and this may also yield different results in the personality trait attribution. Indeed, we chose a monotonous and cyclic behaviour, which may not allow for important motion cues, such as acceleration or pause. These factors have been shown to impact on perceived emotion in dance (e.g., Brownlow, et al., 1997), a display type that further shows potential for communicating individual differences in personality traits. For instance, females rate the dance quality higher when the dancers are conscientious and agreeable males (Fink, et al., 2012).

It could be that other measurements of personality would provide support for the kernel of truth hypothesis. For instance, instead of scores from personality questionnaires, judgements made by known peers, spouses or expert judges could be used. This is treated further under Limitations.

Implications

The consistent consensus in personality trait attribution indicates that there is indeed something that remains in form-degraded body-motion stimuli which influences personality impressions. The motion analyses that we conducted identified dynamic cues that predicted such personality trait attribution. This shows that whole-body motion contains cues not only for gender, identity and emotion, but also for personality. Due to the availability of dynamic body cues in the absence of other cues, this can have implications for human interactions, because we know that trait impressions are automatic and can lead to important social decisions.

A follow-up study (Chapter VI) further showed that motion parameters can be manipulated in order to change perceived personality, which has implications for animators who are interested in creating believable agents for entertainment, training or marketing purposes. Indeed, many blockbusters these days are based entirely on animations, and the audience expects more and more realistic characters in movies or games. Animated TV-commercials are increasingly popular, and it may have massive consequences to the effectiveness of such a commercial if the viewers perceive the character as trustworthy.

The findings pave the way for studies exploring whether motion cues also have an impact on personality attribution in face-to-face interactions or other situations where a target is observed by an audience, such as during sales pitches or speeches. Further studies would need to be carried out to verify this and to assess whether deliberate changes in motion cues can have consequences for trait impressions in real-life situations.

Synthesis

Even if this study has shown consensus amongst observers in their trait impressions based on body-motion, this does not automatically mean that body motion cues are used when other channels, such as the face or voice, are present. However, static information about body size and shape was visible in the full-light displays and the finding that motion parameters were associated with personality judgements for these stimuli thus suggests that body-motion cues can be effective for personality attribution when presented with other cues. Strong correlations were also found between observers’ judgements for the full-light and patch-light versions of the same video sequences, adding further support to the claim that dynamic cues are important even when static cues are visible.

The literature further contains plenty of evidence pointing towards motion cues being able to influence social perception also when other cues are available. First, it has been shown that processing of biological motion is automatic (Thornton & Vuong, 2004); these authors showed that observers are not able to ignore task-irrelevant point-light walkers presented in the periphery when performing the simple task of judging the direction of a central target point-light walker. Second, Wang and Jiang (2012) have also shown that scrambled biological motion sequences of point-light walkers attract more attention than otherwise identical sequences without the biological characteristics (motion acceleration and motion phase). (For additional findings related to the automaticity of biological motion processing, see also Watanabe, 2008.) Finally, it has been shown that social decisions such as voting outcomes can be traced to the candidate's body motion (Kramer, et al., 2010), which suggests that motion processing affects person perception also for more cognitively demanding tasks using audio-visual recordings of humans.

On a related note, research suggests that there is no clear bias towards either the face or the body as cues for emotion processing. Using binocular rivalry in which faces and bodies were presented statically at the same time, Stienen and de Gelder (2011) found that whichever stimulus expressed anger had perceptual dominance over the other. This fits with the anger-superiority found by Chouchourelou and Shiffrar (2006), who showed that observers were more sensitive to point-light walkers where targets were expressing anger. Other studies suggest that which nonverbal channel is the most effective for communicating emotion depends on the emotion expressed; for example, in one study, the body was found to be the preferred—and most effective—channel for communicating pride and the face for anger (App, McIntosh, Reed, & Hertenstein, 2011). Yet other studies suggest that bodily and facial cues interact for communicating emotion: emotionally congruent body-face pairs (e.g., angry body with angry face) were perceived as more emotional, on the emotion dimension in question (e.g., anger), than incongruent body-face pairs (e.g., angry body with fearful face; App, Reed, & McIntosh, 2011).

Interaction of different modalities in person perception has further been found for other judgements of emotion. For instance, if listeners are presented with a face together with a voice, they process the voice in accordance with the gender of the target face (K. Johnson, Strand, & D'Imperio, 1999). But auditory cues can also change visual perception: it has been shown that perceived gender of a point-light walker can be changed using auditory signals (van der Zwan, et al., 2009). In van der Zwan and colleagues' study, gender-ambiguous walkers were judged to be female if they were presented together with sounds of a female auditory walking sequence. These findings further add to the proposition that dynamic cues from whole-body motion may have an impact on personality trait attribution when other cues are also available to observers.

Limitations

Motion capture procedure. A possible limitation of the present work is that the recording was done under highly controlled settings, which may have had an impact on the targets' gait. The lack of a correlation between self-reported personality scores and any of the trait ratings may for instance be due to the instructions given to the target walkers. It has been shown that gender recognition is more accurate when observers are not conscious about gender-typical movements showing natural actions (Runeson & Frykholm, 1983) and, analogously, it could be that the instruction in our study to ‘walk naturally’ lead to an unnatural walk. Furthermore, the target individuals in our study were aware that they were part of a study that was investigating the link between body motion and personality, which may have influenced their behaviour. The targets’ knowledge that they were under observation and that their motion would be scrutinised may have further affected the naturalness of their gait. A follow-up study would benefit from using gait stimuli obtained in naturally occurring situations, wherein the target is unaware that they are being observed. Future advances in motion capture techniques may allow for recordings to be made without preparing a target (i.e. where there is no need to fit retro-reflective markers or body suits) or bring them into prepared motion capture areas.

Motion analyses. As noted, we found no link between self-reported personality and body motion; however, it is possible that other motion analyses may have detected such a link. Indeed, other means of extracting motion parameters from gait are possible, such as change in joint angles (Roether, et al., 2009) or acceleration of specific body parts (Pollick, Paterson, et al., 2001), and it could be that alternative motion-analysis approaches may show a link between a person’s personality and his or her gait.

Although PCA can capture the large amount of motion-capture data (Troje, 2002), it can be criticised for removing redundancy. We used only those components that explained the most amount of variance, and subsequently fitted these components with parameters in such a way that minor deviations were lost. However, the ‘error’ from the sinusoidal fitting could be key to the individual differences in gait, not to mention the potential source of personality cues contained in the 6% of the variance that was not explained by the two chosen components in the first place.

As an example, in the motion data there is a lot of redundancy due to the symmetry of a person's walk. As one arm swings forward the other swings back and so therefore variables involved in the symmetric motion of the two arms will be grouped together in the PCA. However, it could be that diversion away from this symmetry can influence trait ratings. Plenty of evidence point towards the importance of symmetry to a person's health and reproductive success (Gangestad & Simpson, 2000; B. C. Jones, et al., 2001; Møller & Thornhill, 1998), and dynamic symmetry has been found to affect emotional expressiveness in faces (Penton-Voak & Chang, 2008) and whole-body movement (Roether, et al., 2008). Cutting and Proffitt (1981) indicated more than three decades ago that symmetry is key to gait quality. Jerkiness is also linked to perceived energy from arm movement (Pollick, Paterson, et al., 2001), and although scaling our motion parameters may impact on jerkiness, they are not likely in themselves to reflect this motion cue, which may have been lost through the sinusoidal fitting procedure. We can therefore not rule out that other motion analyses are able to detect individual differences of personality in gait.

It would be interesting in the future to determine whether the residuals from the modelling procedure contain information about individual personality. However, there is currently no straightforward means to analyse such data.

Personality measurements. Similar to the issue with the choice of motion analysis, we have not yet addressed the possibility that other ways of measuring personality might be superior to the one we have opted for. To begin with, the NEO FFI, SF by Costa and McCrae (1992) is not the only personality questionnaire;

others include the six-factor Hexaco personality questionnaire (Ashton & Lee, 2009)—which adds the scale honesty-humility to the Big Five—or Eysenck and Eysenck's (1975) three-factor personality questionnaire, which consists of the factors Extraversion, Neuroticism and Psychoticism. The NEO PI questionnaire was chosen because it has been repeatedly validated in the literature (Costa & McCrae, 1992; Costa, et al., 2001; McCrae & Costa, 1985, 1987) and it has strong correlations with external criteria (see Chapter I). However, follow-up studies may consider using alternative questionnaires for assessing personality.

However, self-reports are not the only way of assessing someone's personality, and some argue that they are not the best procedure for doing so. In fact, the strongest opponents of the use of self-reports as a valid source for information about personality might argue that the only way to describe personality, 'if there is such a thing,' is through judgements made by others (Funder, 1995; Swann, 1984). Other, less stringent, researchers argue that peer-reports are better estimators of external validity criteria, such as job performance (Oh, et al., 2011; Zimmerman, et al., 2010). Finally, the Self-other Knowledge Asymmetry model (Vazire, 2010) suggests that individuals know themselves best on some traits (e.g., neuroticism) whilst others are better judges on traits that may have observable criteria (e.g., intellect). Scores on the five-factor inventory derived from ratings by knowledgeable others are found to be strong predictors of work performance (Oh, et al., 2011) and some research suggest that a peer-self agreement exists only for subsets of the factors; e.g., extraversion (Vazire, 2010).

It would thus be interesting to assess whether our observers' personality judgements correspond to personality traits obtained by known peers, or expert judges. A modification of this study could also expose observers to increasingly complex stimuli, perhaps video clips with sound, or even letting observers and targets interact for varying amounts of time, in order to see whether judgements converge across these conditions. Additionally, trait ratings could be correlated with other aspects of the target individuals lives, such as success in relationships, job performance or academic achievements.

Traits and labels. The use of traits can be criticised at a general level. For instance, Mischel (1968) pointed out the dangers of tautologous reasoning when inferring traits from behaviour, in order to explain this same behaviour. For instance, people are quick at deducing that someone is anxious, if they observe this person behaving ‘anxiously’. However, even if it is the case that traits are simply reputations—labels defined by observers—this does not mean they are not central to people’s lives. Firstly, reputations may influence behaviour: most people are concerned with their reputations, sometimes to an extreme, such as committing suicide because of ‘lost face’ (Hogan, 1996). Secondly, reputations will affect how you are treated; the fact that letters of recommendation are still used as a preliminary decision to invite job applicants for interviews is testament to this.

Leaving criticism of the use of traits aside, we still have not tackled the potential issue arising from the choice of trait scales in this thesis. The antonyms used will affect the decisions by observers and the chosen labels are likely to have an impact on the conclusions drawn from the trait ratings. We used trait terms from Heberlein, et al. (2004), and also verified the loading of the trait terms on their respective trait constructs in the validation study by McCrae and Costa (1987). The factor analyses reported in Chapters IV and VIII also confirmed that the trait scales loaded on their anticipated component (i.e. either dominance or valence; see Fiske, et al., 2007). Thus, although we are open to the possibility that other trait adjectives may have yielded results somewhat diverging from those reported in this thesis, we stand behind our choice of trait scales.

Individual differences among observers. This thesis has not explored individual differences in the observers and how these may affect trait attribution. Large individual differences exist in the ability to accurately detect gender from body-motion (Pollick, et al., 2005). Autistic traits within a healthy population have also been shown to be negatively linked to sensitivity to anger displays in point-light walkers (Kaiser & Shiffrar, 2008). Investigating outliers among our personality trait judges may therefore lead to an improvement in the methodology.

Male and female observers also differ: it has been shown that females are better than men at distinguishing biological from non-biological motion, as well as recognising emotions from point-light walkers (Alaerts, et al., 2011). However, studies of other forms of person perception have not reported gender differences; for instance, physical attractiveness (Berscheid, et al., 1971). Nonetheless, future studies may benefit from controlling for observer gender.

We examined a link between personality scores from the NEO FFI SF and these observers’ trait judgements, but found no such association (see Appendix). Because no link was identified after the first five experiments, we decided to stop collecting personality traits in order to reduce the burden on participants. However, we do not reject the possibility that future studies may find such links, and it may be beneficial to collect data on other differences, such as cultural background.

Future Directions

We have shown that instructing targets to display different actions or emotions can influence personality trait judgements of that target and, using point-light displays of whole-body motion, we have shown that personality trait impressions can be changed by manipulating dynamic aspects of the gait of a target. This paves the way for investigating whether targets can be instructed to move in ways that affect how they are perceived. In the first instance, one could give different motion-related instructions to targets during the motion capture procedures (e.g., ‘use more of the space around you’), and see whether observer ratings based on point-light walkers can be affected by these instructions. Different emotions or actions can be investigated separately. Secondly, such instructions can be used for increasingly rich stimuli, including real-life interactions such as job interviews, political speeches, or sales pitches. One could test, for example, whether exaggerated use of personal space will lead to people appearing more positive, or whether this dynamic feature is tied to something intrinsic to the target that cannot successfully be modified.

For creators of avatars, it would certainly be of interest to investigate whether scalable motion cues can be identified for actions other than walking. Interacting with objects, e.g., moving a chair, will not necessarily see the same cyclic patterns as walking, and it is likely that different motion analyses would be required. However, the results presented in this thesis are highly suggestive that it is indeed possible to extract motion cues also for other types of action, and it should be possible to manipulate these motion cues also for other displayed behaviours. Follow-up studies should include other display formats, from stick-figures to full-body displays complete with superficial details such as skin colour and clothing.

Trait attribution has in this thesis been studied using explicit trait impressions. Whilst it has been shown that explicit ratings are reflective of implicit impressions given their external criteria of election outcome (Little, et al., 2007) and student module feedback (Ambady & Rosenthal, 1993), for example, follow-up studies could benefit from using paradigms assessing implicit judgements. For instance, the stimuli can be used in economic games, cueing paradigms, or in incidental learning paradigms.

General Conclusion

This thesis has showed that people use body-motion to make personality trait judgements of strangers. Although we cannot conclude categorically that there is no link between the personality traits of a target walker and how that person is perceived by an observer, it is clear from findings that people attribute personality traits perhaps too readily.

It is not the wish of the authors that knowledge obtained from our work be misused for manipulating people. Rather, we stress that one should be aware of biases and tendencies in overgeneralisations in trait attribution. That is, whenever you meet someone you should consider that you are likely to be judged by the way in which you move your body. Similarly, you are likely to unconsciously form your own impressions of the person you meet based on these same cues. Awareness of these biases may be important, considering how they are often incorrect, and considering the huge implications that first impressions may have on many aspects of human life.

Appendix

Appendix: Correlations (ρ) between Self-Reported Personality Scores and Trait Ratings by Observers from Four Experiments

	Self-reported personality trait				
Trait scale	N	E	O	A	C
Approach.	-.28**	.05	.18	.00	.06
Extraversion	.20	-.04	-.08	-.05	-.08
Neuroticism	.21*	-.04	-.07	-.05	-.14
NP	.13	-.07	-.06	.01	-.10
Trustworth.	-.06	-.14	.03	-.03	.02
Warmth	-.21*	-.06	.21*	-.02	.04

Note. Approach. = approachability, NP = novelty-preference, Trustworth. = trustworthiness, N = neuroticism, E = extraversion, O = openness to experience, A = agreeableness, C = conscientiousness. ***** $p < .05$; ****** $p < .01$. Data are from 92 observers from Experiments 1.1, 1.2, 1.3 (Chapter II), and 2.1 (Chapter III) who were asked to fill out a paper-version of the NEO FFI questionnaire immediately upon completing the trait rating. Once corrected for multiple comparisons ($n = 30$), there were no significant correlation coefficients between any of the NEO components and the trait scales.

References

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature, 393*(6684), 470-474.

Alaerts, K., Nackaerts, E., Meyns, P., Swinnen, S. P., & Wenderoth, N. (2011). Action and emotion recognition from point light displays: An investigation of gender differences. *PLoS ONE, 6*(6).

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, *55*(3), 387-395.

Albright, L., Malloy, T. E., Dong, Q., Kenny, D. A., Fang, X. Y., & Winquist, L. (1997). Cross-cultural consensus in personality judgments. *Journal of Personality and Social Psychology*, *72*(3), 558-569.

Allik, J., Realo, A., Mõttus, R., Borkenau, P., Kuppens, P., & Hřebíčková, M. (2010). How people see others is different from how people see themselves: A replicable pattern across cultures. *Journal of Personality and Social Psychology, 99*(5), 870-882.

Allport, G. W. (1937). *Personality: a psychological interpretation*. Oxford England: Holt.

Allport, G. W. (1961). *Pattern and growth in personality*. Oxford England: Holt, Reinhart & Winston.

Allport, G. W., & Odbert, H. S. (1936). Trait-Names A Psycho-lexical Study. *Psychological Monographs, 47*(1), 1-171.

Allport, G. W., & Postman, L. (1947). *The psychology of rumor*. Oxford England: Henry Holt.

Amaya, K., Bruderlin, A., & Calvert, T. (1996). Emotion from motion. In W.A Davis & R Bartels (Eds.), *Graphics Interface* (pp. 228-229). Toronto, Canada: Canadian Human-Computer Communications Society.

Ambady, N., & Rosenthal, R. (1993). Half a minute - Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, *64*(3), 431-441.

Anderson, C., John, O. P., Keltner, D., & Kring, A. M. (2001). Who attains social status? Effects of personality and physical attractiveness in social groups. *Journal of Personality and Social Psychology, 81*(1), 116-132.

App, B., McIntosh, D. N., Reed, C. L., & Hertenstein, M. J. (2011). Nonverbal Channel Use in Communication of Emotion: How May Depend on Why. *Emotion, 11*(3), 603-617.

App, B., Reed, C. L., & McIntosh, D. N. (2011). Relative contributions of face and body configurations: Perceiving emotional state and motion intention. *Cognition & Emotion, 26*(4), 690-698.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*(3), 258-290.

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: a short measure of the major dimensions of personality. *Journal of Personality Assessment, 91*(4), 340-345.

Atkinson, A. P. (2009). Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders. *Neuropsychologia, 47*(13), 3023-3029.

Atkinson, A. P., & Adolphs, R. (2005). Visual emotion perception: Mechanisms and processes. In L. F. Barrett, P. M. Niedenthal & P. Winkielman (Eds.), *Emotion and consciousness* (pp. 150-182). New York, NY US: Guilford Press.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception, 33*(6), 717-746.

Atkinson, A. P., Heberlein, A. S., & Adolphs, R. (2007). Spared ability to recognise fear from static and moving whole-body cues following bilateral amygdala damage. *Neuropsychologia, 45*(12), 2772-2782.

Atkinson, A. P., Tunstall, M. L., & Dittrich, W. H. (2007). Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. *Cognition, 104*(1), 59-72.

Bänziger, T., Grandjean, D., & Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: The Multimodal Emotion Recognition Test (MERT). *Emotion*, *9*(5), 691-704.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6*(2), 269-278.

Barrett, H. C., Todd, P. M., Miller, G. F., & Blythe, P. W. (2005). Accurate judgments of intention from motion cues alone: A cross-cultural study. *Evolution and Human Behavior, 26*(4), 313-331.

Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology, 83*(3), 377-391.

Berry, D. S. (1990). Taking people at face value: Evidence for the Kernel of Truth Hypothesis. *Social Cognition, 8*, 883-893.

Berscheid, E., Dion, K., Walster, E., & Walster, G. W. (1971). Physical attractiveness and dating choice: A test of the matching hypothesis. *Journal of Experimental Social Psychology, 7*(2), 173-189.

Bond Jr, C. F., Berry, D. S., & Omar, A. (1994). The Kernel of Truth in judgments of deceptiveness. *Basic & Applied Social Psychology, 15*, 523-534.

Bono, J. E., & Vey, M. A. (2007). Personality and emotional performance: Extraversion, neuroticism, and self-monitoring. *Journal of Occupational Health Psychology, 12*(2), 177-192.

Boone, R. T., & Cunningham, J. G. (1998). Children's decoding of emotion in expressive body movement: The development of cue attunement. *Developmental Psychology, 34*(5), 1007-1016.

Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at Zero Acquaintance. *Journal of Personality and Social Psychology, 62*(4), 645-657.

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, *86*(4), 599-614.

Brownlow, S., Dixon, A. R., Egbert, C. A., & Radcliffe, R. D. (1997). Perception of movement and dancer characteristics from point-light displays of dance. *The Psychological Record, 47*(3), 411-421.

Bruderlin, A., & Williams, L. (1995). Motion signal processing. Paper presented at the Proceedings of the 22nd annual conference on Computer graphics and interactive techniques.

Calvo-Merino, B., Jola, C., Glaser, D. E., & Haggard, P. (2008). Towards a sensorimotor aesthetics of performing art. *Consciousness and Cognition, 17*(3), 911-922.

Camurri, A., De Poli, G., Leman, M., & Volpe, G. (2001). *A multi-layered conceptual framework for expressive gesture applications*. Paper presented at the Proceedings of the workshop of on current research directions in computer music.

Camurri, A., Lagerlof, I., & Volpe, G. (2003). Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies, 59*(1-2), 213-225.

Camurri, A., Mazzarino, B., Trocca, R., & Volpe, G. (2002). Real-time analysis of expressive cues in human movement: University of Michigan, Ann Arbor: International Computer Music Conference.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25*(2), 215-224.

Chouchourelou, A., Matsuka, T., Harber, K., & Shiffrar, M. (2006). The visual analysis of emotional actions. *Social Neuroscience, 1*(1), 63-74.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments & Computers, 25*(2), 257-271.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (2007). PsyScope (version X. Build 51) [Computer Software] (Version Version X, build 51). Pittsburgh, Pennsylvania: Carnegie Mellon University. Retrieved April 02, 2009. Available from http://psy.ck.sissa.it.

Colvin, C. R., & Funder, D. C. (1991). Predicting personality and behavior: A boundary on the acquaintanceship effect. *Journal of Personality and Social Psychology, 60*(6), 884-894.

Costa, P. T., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology, 54*(5), 853-863.

Costa, P. T., & McCrae, R. R. (1992). NEO P1-R. Professional manual. Odessa, FL: Psychological Assessment Resources, Inc.

Costa, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, *81*(2), 322-331.

Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, *28*(2), 117-139.

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society, 9*(5), 353-356.

Cutting, J. E., & Proffitt, D. R. (1981). Gait perception as an example of how we may perceive events. In R. Walk & H. L. Pick (Eds.), *Intersensory perception and sensory integration* (pp. 249-273). New York: Plenum.

Dekeyser, M., Verfaillie, K., & Vanrie, J. (2002). Creating stimuli for the study of biological-motion perception. *Behavior Research Methods*, *Instruments & Computers*, *34*(3), 375-382.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*(6), 1246-1256.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*(3), 285-&.

Dittrich, W. H., Troscianko, T., Lea, S. E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception, 25*(6), 727-738.

Ekman, P. (2003). Darwin, deception, and facial expression. *Annals of the New York Academy of Sciences, 1000*, 205-221.

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: Automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience, 19*(9), 1508-1519.

Eysenck, H. J. (1975). *Manual of the Eysenck personality questionnaire (Junior & adult)*. Hodder and Stoughton Educational: Sevenoaks.

Fan, J., Dai, W., Liu, F., & Wu, J. (2005). Visual perception of male body attractiveness. *Proceedings of the Royal Society: Biological Sciences, 272*(1560), 219-226.

Fan, J., Liu, F., Wu, J., & Dai, W. (2004). Visual perception of female physical attractiveness. *Proceedings of the Royal Society: Biological Sciences, 271*(1537), 347-352.

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin, 111*(2), 304.

Fiedler, K., & Schenck, W. (2001). Spontaneous Inferences From Pictorially Presented Behaviors. *Personality and Social Psychology Bulletin, 27*(11), 1533-1547.

Fiedler, K., Schenck, W., Watling, M., & Menges, J. I. (2005). Priming Trait Inferences Through Pictures and Moving Pictures: The Impact of Open and Closed Mindsets. *Journal of Personality and Social Psychology, 88*(2), 229-244.

Fink, B., Weege, B., Flügge, J., Röder, S., Neave, N., & McCarty, K. (2012). Men's personality and women's perception of their dance quality. *Personality and Individual Differences, 52*(2), 232-235.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77-83.

Friedman, H. S., DiMatteo, M. R., & Taranta, A. (1980). A study of the relationship between individual differences in nonverbal expressiveness and factors of personality and social interaction. *Journal of Research in Personality, 14*(3), 351-364.

Frigerio, E., Burt, D. M., Gagliardi, C., Martelli, S., Borgatti, R., Perrett, D. I., et al. (2006). Is everybody always my friend? Perception of approachability in Williams syndrome. *Neuropsychologia, 44*(2), 254-259.

Funder, D. C. (1995). On the accuracy of personality judgment: a realistic approach. *Psychological Review, 102*(4), 652-670.

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55*(1), 149-158.

Galton, F. (1884). The Measurement of Character. *Fortnightly Review, 36*, 179-185.

Gangestad, S. W., & Simpson, J. A. (2000). The evolution of human mating: Trade-offs and strategic pluralism. *Behavioral and Brain Sciences, 23*(4), 573-644.

Goldberg, L. R. (1990). An alternative description of personality - the Big-5 Factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216-1229.

Gross, M., Crane, E., & Fredrickson, B. (2010). Methodology for assessing bodily expression of emotion. *Journal of Nonverbal Behavior, 34*(4), 223-248.

Gunns, R. E., Johnston, L., & Hudson, S. M. (2002). Victim selection and kinematics: A point-light investigation of vulnerability to attack. *Journal of Nonverbal Behavior, 26*(3), 129-158.

Hager, J. C., & Ekman, P. (1979). Long-distance of transmission of facial affect signals. *Ethology and Sociobiology, 1*(1), 77-82.

Heberlein, A. S., Adolphs, R., Tranel, D., & Damasio, H. (2004). Cortical regions for judgments of emotions and personality traits from point-light walkers. *Journal of Cognitive Neuroscience, 16*(7), 1143-1158.

Heberlein, A. S., & Saxe, R. R. (2005). Dissociation between emotion and personality judgments: Convergent evidence from functional neuroimaging. *NeuroImage, 28*(4), 770-777.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology, 57,* 243-259.

Hill, H., & Pollick, F. E. (2000). Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychological Science, 11*(3), 223-228.

Hogan, R. (1996). A socioanalytic perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives.* (pp. 163-179). New York, NY US: Guilford Press.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14*(2), 201-211.

Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics, 27*(4), 359-384.

Johnson, K. L., McKay, L. S., & Pollick, F. E. (2011). He throws like a girl (but only when he's sad): emotion affects sex-decoding of biological motion displays. *Cognition, 119*(2), 265-280.

Johnston, L., Hudson, S. M., Richardson, M. J., Gunns, R. E., & Garner, M. (2004). Changing kinematics as a means of reducing vulnerability to physical attack. *Journal of Applied Social Psychology, 34*(3), 514-537.

Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2004). When facial attractiveness is only skin deep. *Perception, 33*(5), 569-576.

Jones, B. C., Little, A. C., Penton-Voak, I. S., Tiddeman, B. P., Burt, D. M., & Perrett, D. I. (2001). Facial symmetry and judgements of apparent health - Support for a "good genes" explanation of the attractiveness-symmetry relationship. *Evolution and Human Behavior, 22*(6), 417-429.

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist, 34*(2), 107-117.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology, 3*(1), 1-24.

Kaiser, M. D., & Shiffrar, M. (2008). Individual differences in perceptual sensitivity to emotional human movement. *Visual Cognition, 16*(8), 1115-1119.

Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception - Acquaintance and the Big-5. *Psychological Bulletin, 116*(2), 245-258.

Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L.-c. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology, 62*(1), 88-97.

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior, 20*(3), 165-182.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics, 21*(6), 575-580.

Kozlowski, L. T., & Cutting, J. E. (1978). Recognizing the gender of walkers from point-lights mounted on ankles: Some second thoughts. *Attention, Perception, & Psychophysics, 23*(5), 459-459.

Kramer, R. S. S., Arend, I., & Ward, R. (2010). Perceived health from biological motion predicts voting behaviour. *Quarterly Journal of Experimental Psychology, 63*(4), 625-632.

Kring, A. M., & Sloan, D. M. (2007). The Facial Expression Coding System (FACES): Development, validation, and utility. *Psychological Assessment, 19*(2), 210-224.

Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007). Facial appearance affects voting decisions. *Evolution and Human Behavior, 28*(1), 18-27.

Loehlin, J. C., & Martin, N. G. (2001). Age changes in personality traits and their heritabilities during the adult years: Evidence from Australian twin registry samples. *Personality and Individual Differences, 30*(7), 1147-1160.

Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance, 31*(1), 210-220.

McDonnell, R., Jörg, S., McHugh, J., Newell, F. N., & O'Sullivan, C. (2009). Investigating the role of body shape on the perception of emotion. ACM Transactions on Applied Perception, 6(3), 1-11.

Mather, G., & Murdoch, L. (1994). Gender Discrimination in Biological Motion Displays Based on Dynamic Cues. *Proceedings of the Royal Society of London: Biological Sciences, 258*(1353), 273-279.

Mattes, K., Spezio, M., Kim, H., Todorov, A., Adolphs, R., & Alvarez, R. M. (2009). Predicting Election Outcomes from Positive and Negative Trait Assessments of Candidate Images. *Political Psychology, 31*(1), 41-58.

McCrae, R. R., & Costa, P. T. (1985). Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Personality and Individual Differences, 6*(5), 587-597.

McCrae, R. R., & Costa, P. T. (1987). Validation of the 5-Factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81-90.

McCrae, R. R., & Costa, P. T. (1989). Rotation to maximize the construct validity of factors in the NEO Personality Inventory. *Multivariate Behavioral Research, 24*(1), 107-124.

McCrae, R. R., Costa, P. T., & Wiggins, J. S. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model *The five-factor model of personality: Theoretical perspectives.* (pp. 51-87). New York, NY US: Guilford Press.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*(2), 175-215.

Mischel, W. (1968). *Personality and assessment*. Hoboken, NJ US: John Wiley & Sons Inc.

Møller, A. P., & Thornhill, R. (1998). Bilateral Symmetry and Sexual Selection: A Meta-Analysis. *The American Naturalist, 151*(2), 174-192.

Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior, 27*(4), 237-254.

Montepare, J. M., & Zebrowitz-McArthur, L. (1988). Impressions of people created by age-related qualities of their gaits. *Journal of Personality and Social Psychology, 55*(4), 547-556.

Noor, F., & Evans, D. C. (2003). The effect of facial symmetry on perceptions of personality and attractiveness. *Journal of Research in Personality, 37*(4), 339-347.

Oh, I.-S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology, 96*(4), 762-773.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 105*(32), 11087-11092.

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion, 9*(1), 128-133.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin, 49*(3), 197-237.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Oxford, England: Univer. Illinois Press.

Parkinson, B. (2005). Do Facial Movements Express Emotions or Communicate Motives? *Personality and Social Psychology Review, 9*(4), 278-311.

Penton-Voak, I. S. (2009). Personal communication.

Penton-Voak, I. S., & Chang, H. Y. (2008). Attractiveness judgements of individuals vary across emotional expression and movement conditions. *Journal of Evolutionary Psychology, 6*(2), 89-100.

Penton-Voak, I. S., Pound, N., Little, A. C., & Perrett, D. I. (2006). Personality judgments from natural and composite facial images: More evidence for a "kernel of truth" in social perception. *Social Cognition, 24*(5), 607-640.

Polinko, N. K., & Popovich, P. M. (2001). Evil thoughts but angelic actions: Responses to overweight job applicants. *Journal of Applied Social Psychology, 31*(5), 905-924.

Pollick, F. E., Fidopiastis, C., & Braden, V. (2001). Recognising the style of spatially exaggerated tennis serves. *Perception, 30*(3), 323-338.

Pollick, F. E., Kay, J. W., Heim, K., & Stringer, R. (2005). Gender recognition from point-light walkers. *Journal of Experimental Psychology-Human Perception and Performance, 31*(6), 1247-1265.

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition, 82*(2), B51-BB61.

Porter, M. A., Coltheart, M., & Langdon, R. (2007). The neuropsychological basis of hypersociability in Williams and Down syndrome. *Neuropsychologia, 45*(12), 2839-2849.

Prasad, S., & Shiffrar, M. (2009). Viewpoint and the recognition of people from their movements. *Journal of Experimental Psychology-Human Perception and Performance, 35*(1), 39-49.

Puhl, R., & Brownell, K. D. (2001). Bias, Discrimination, and Obesity. *Obesity, 9*(12), 788-805.

Riggio, R. E., & Friedman, H. S. (1986). Impression formation: The role of expressive behavior. *Journal of Personality and Social Psychology, 50*(2), 421-427.

Riggio, R. E., Lippa, R., & Salinas, C. (1990). The display of personality in expressive movement. *Journal of Research in Personality, 24*(1), 16-31.

Roether, C. L., Omlor, L., Christensen, A., & Giese, M. A. (2009). Critical features for the perception of emotion from gait. *Journal of Vision, 9*(6), 1-32.

Roether, C. L., Omlor, L., & Giese, M. A. (2008). Lateral asymmetry of bodily emotion expression. [Letter]. *Current Biology, 18*(8), R329-R330.

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A Multidimensional Approach To The Structure Of Personality Impressions. *Journal of Personality and Social Psychology, 9*(4), 283-294.

Rosenthal, R. (1973). Estimating effective reliabilities in studies that employ judges' ratings. *Journal of Clinical Psychology, 29*(3), 342-345.

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General, 112*(4), 585-615.

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion, 9*(2), 260-264.

Sakaguchi, K., & Hasegawa, T. (2006). Person perception through gait information and target choice for sexual advances: Comparison of likely targets in experiments and real life. *Journal of Nonverbal Behavior, 30*(2), 63-85.

Schindler, K., Van Gool, L., & de Gelder, B. (2008). Recognizing emotions expressed by body pose: A biologically inspired neural model. *Neural Networks, 21*(9), 1238-1246.

Smith, F. W., & Schyns, P. G. (2009). Smile through your fear and sadness: transmitting and identifying facial expression signals over a range of viewing distances. *Psychological Science, 20*(10), 1202-1208.

Stienen, B. M. C., & de Gelder, B. (2011). Fear modulates visual awareness similarly for facial and bodily expressions. *Frontiers in Human Neuroscience, 5*.

Swami, V., Einon, D., & Furnham, A. (2006). The leg-to-body ratio as a human aesthetic criterion. *Body Image, 3*(4), 317-323.

Swami, V., & Tovee, M. J. (2005). Female physical attractiveness in Britain and Malaysia: a cross-cultural study. *Body Image, 2*(2), 115-128.

Swann, W. B. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review, 91*(4), 457-477.

Thornton, I. M., & Vuong, Q. C. (2004). Incidental Processing of Biological Motion. *Current Biology, 14*(12), 1084-1089.

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social Cognitive and Affective Neuroscience, 3*(2), 119-127.

Todorov, A., & Duchaine, B. (2008). Reading trustworthiness in faces without recognizing faces. *Cognitive Neuropsychology, 25*(3), 395-410.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308*(5728), 1623-1626.

Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision, 2*(5), 371-387.

Troje, N. F., & Szabo, S. (2006). Why is the average walker male? *Journal of Vision, 6*(6), 1034.

Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics, 67*(4), 667-675.

Uleman, J. S. (2005). On the inherent ambiguity of traits and other mental concepts. In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 253-267). New York: Guilford Publications.

Uleman, J. S., Blader, S., & Todorov, A. (2005). Implicit impressions. In R. R. Hassin, J. S. Uleman & J. A. Bargh (Eds.), *The new unconscious* (pp. 362-392). New York: Oxford University Press.

Uleman, J. S., & Moskowitz, G. B. (1994). Unintended effects of goals on unintended inferences. *Journal of Personality and Social Psychology, 66*(3), 490-501.

van't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition, 108*(3), 796-803.

van der Zwan, R., Machatch, C., Kozlowski, D., Troje, N. F., Blanke, O., & Brooks, A. (2009). Gender bending: auditory cues affect visual judgements of gender in biological motion displays. *Experimental Brain Research, 198*(2-3), 373-382.

Vanrie, J., Dekeyser, M., & Verfaillie, K. (Writer) (2004). Bistability and biasing effects in the perception of ambiguous point-light walkers, *Perception*. United Kingdom: Pion.

Vanrie, J., & Verfaillie, K. (2004). Perception of biological motion: A stimulus set of human point-light actions. *Behavior Research Methods, 36*(4), 625-629.

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology, 98*(2), 281-300.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology, 28*(6), 879-896.

Wang, L., & Jiang, Y. (2012). Life motion signals lengthen perceived temporal duration. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 109*(11), e673-e677.

Watanabe, K. (2008). Behavioral speed contagion: Automatic modulation of movement timing by observation of body movements. *Cognition, 106*(3), 1514-1524.

Williams, J. E., Satterwhite, R. C., & Best, D. L. (1999). Pancultural gender stereotypes revisited: The Five Factor Model. *Sex Roles, 40*(7-8), 513-525.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592-598.

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology, 47*(2), 237-252.

Zimmerman, R. D., Triana, M. d. C., & Barrick, M. R. (2010). Predictive criterion-related validity of observer ratings of personality and job-related competencies using multiple raters and multiple performance criteria. *Human Performance, 23*(4), 361-378.

Addendum

The experiments presented in Chapters V, VI and VII have been accepted for publication:

Thoresen, J. C., Vuong, Q. C., & Atkinson, A. P. (2012). First impressions: Gait cues drive reliable trait judgements. *Cognition*, *124*(3), 261-271.